Clostridium difficile Mannan and *Campylobacter jejuni* Heptan

by

Pawel Gabryelski

A Thesis
presented to
The University of Guelph

In partial fulfillment of requirements
for the degree of
Master of Science
In
Chemistry

Guelph, Ontario, Canada

Abstract

Clostridium difficile Mannan and *Campylobacter jejuni* Heptan

Pawel Gabryelski **Advisor:**

University of Guelph 2015 **Professor Mario A. Monteiro**

Clostridium difficile infection in humans is on the rise and vaccine targets that will reduce both disease and colonization levels are needed. Potential vaccines comprised of *C. difficile* cell-wall polysaccharides PS-I, PS-II, and PS-III are being investigated in the Monteiro laboratory at Guelph. In this work, the discovery of mannan moieties in *C. difficile* is reported. The mannose units were present as end-groups [α-Man-(1→], 2-linked linear units [→2)-α-Man-(1→] and 2,6-linked branch-points [→2,6)-α-Man-(1→]. Within the limits of detection, no mannose residues were detected in the analysis of the growth media. Sporulating *C. difficile* preparations were observed to contain greater amounts of mannose.

Campylobacter jejuni (C. jejuni) is a food borne pathogen that is one of the leading causes of gastroenteritis in the world. Characterization of the surface polysaccharides is a crucial step in the development of a vaccine against *C. jejuni* infection. Knowledge about polysaccharide structure will also give insight into the function of bacterial surface and aid in conjugate vaccine development. In this study, the capsular polysaccharide of *C. jejuni* serotype HS:44 was investigated. Physico-chemical analysis of HS:44 capsular polysaccharide revealed that it was rich in heptose units.

Acknowledgements

I would like to thank my advisor, Dr. Mario A. Monteiro, for all his help and support and for granting me the opportunity to work in his research group.

I would like to thank the members of my advisory committee: Dr. Paul Rowntree and Dr. Adrian Schwan.

I would like to thank the past and present members of the Monteiro research group: Dr. Zuchao Ma, Dr. Lisa Bertolo, Moez Valliani, Yuening Jiao, Brittany Pequegnat, Eman Omari and Olena Redkyna. Thank you for all your help in the labs.

Most importantly, I want to thank my family and friends for all their help and support.

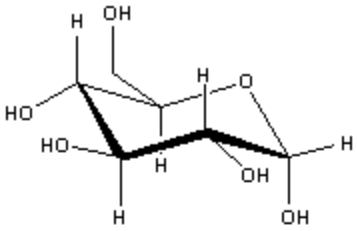
List of Abbreviations

1D	One dimensional
2D	Two dimensional
α	Alpha
β	Beta
δ	Chemical shift
AA	Alditol acetate
APCs	Antigen presenting cells
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
C. difficile	Clostridium difficile
CDAD	*C. difficile*-associated diarrhea
CDI	*C. difficile* infection
C. botilinum	Clostridium botilinum
C. butyricum	Clostridium butyricum
C. jejuni	Campylobacter jejuni
COSY	Correlation Spectroscopy

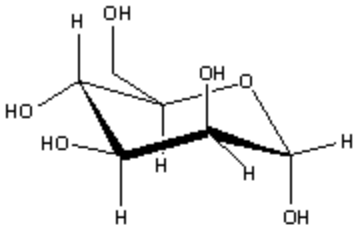
CPS	Capsular Polysaccharide
CRM197	Mutant diphtheria toxoid
DCM	Dichloromethane
dH2O	Deionized water
DT	Diphtheria toxoid
DMSO	Dimethyl sulfoxide
FID	Flame ionization detector
FID	Free induction decay
Gal	Galactose
GC-MS	Gas chromatography mass spectrometry
Glc	Glucose
HMBC	Heteronuclear Multiple Bond Correlation Spectroscopy
HPLC	High performance liquid chromatography
HSQC	Heteronuclear Single Quantum Correlation Spectroscopy
IgG2	Immunoglobulin G2
IgM	Immunoglobulin M
Kdo	Ketodeoxyoctonate

LC-MS	Liquid chromatography mass spectrometry
LOS	Lipo-oligosaccharide
LPS	Lipopolysaccharide
LTA	Lipoteichoic acids
MHC	Major histocompatibility complex
Man	Mannose
MWCO	Molecular weight cut off
NAP1	North American pulsotype 1
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser-Effect Spectroscopy
PMAA	Partially Methylated Alditol Acetate
RF	Radio frequency
Rha	Rhamnose
TOCSY	Total Correlation Spectroscopy
TSP	3-trimethylsilyl-tetradeutero sodium propionate
TT	Tetanus toxoid
WTA	Wall teichoic acids

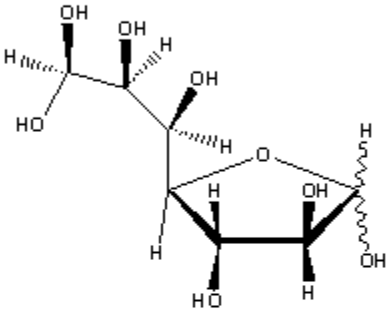
Relevant Structures



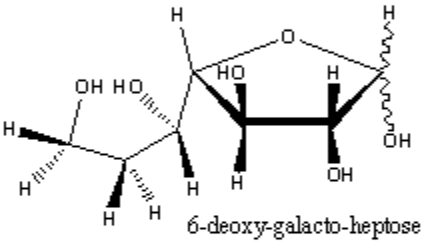
D-Glucose



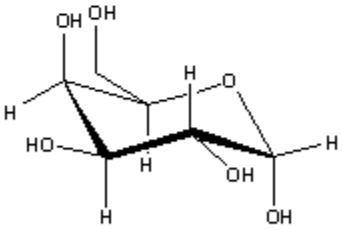
D-Mannose



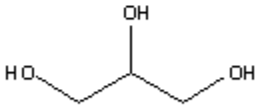
6-deoxy-altro-heptose



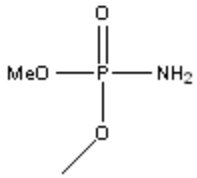
6-deoxy-galacto-heptose



D-Galactose



Glycerol



O-Methyl Phosphoramidate (MeOPN)

Table of Contents

List of Figures

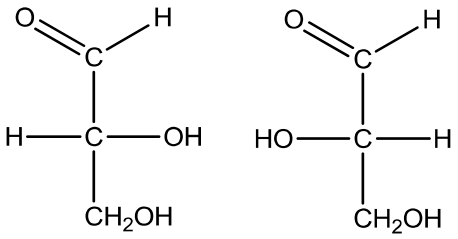
List of Tables

Chapter 1: Introduction

1.1 General Carbohydrates

Carbohydrates are polyhydroxy aldehydes and ketones. They are most commonly found in hemiacetal and acetal forms. Simple carbohydrates are known as sugars or saccharides. Monosaccharides are the simplest carbohydrates as they cannot be hydrolyzed into smaller carbohydrates. Disaccharides and trisaccharides yield two and three monosaccharides. Oligosaccharides yield 2-10 monosaccharides. Polysaccharides are composed of 10 or more monosaccharides [1]

Monosaccharides are classified according to 3 characteristics: the placement of the carbonyl group, the number of carbon atoms in the carbon chain and the structure of the chiral center. If the carbonyl functional group is an aldehyde, the monosaccharide is an aldose; if the carbonyl group is a ketone, the monosaccharide is a ketose. Monosaccharides containing three carbon atoms are called trioses; four carbon atoms is a tetrose; five carbon atoms is a pentose; six carbon atoms is a hexose; and one containing seven carbon atoms is a heptose. The two systems are easily combined. For example, an aldose with 4 carbons is called an aldotetrose. Optical isomers are distinguished by the spatial configuration of its atoms. The D and L nomenclature designation system uses the simplest chiral monosaccharide glyceraldehyde as a configurational standard for all monosaccharides. [1]

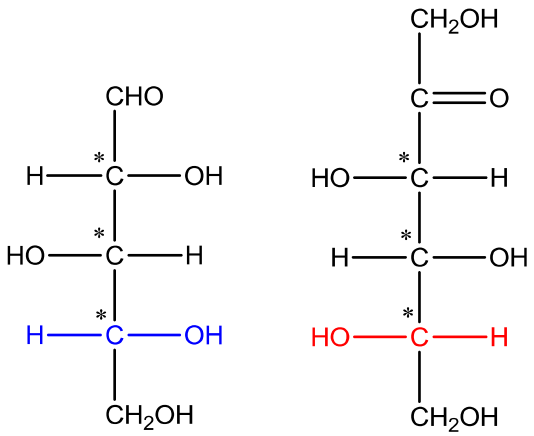


D-Glyceraldehyde L-Glyceraldehyde

Figure 1: D and L Glyceraldehyde

Carbohydrates are depicted in various ways. Each representation gives insight on the conformation, spatial orientation and chirality of the carbohydrate. The Fischer projection displays the carbohydrate in an open chain form. The chain is oriented so that C1 is at the top displaying the lowest chiral center at the bottom of the chain. Therefore, The Fischer projection is useful for quick differentiation between D and L isomers.

Glyceraldehyde contains one chiral center. When the molecule is oriented with the aldehyde group on the top of the molecule and the hydroxyl group is on the right of this chiral center the molecule is designated as D-glyceraldehyde; if the hydroxyl group is on the left the molecule is designated as L-glyceraldehyde. Chirality of monosaccharides is determined from the asymmetric carbon that is furthest from the carbonyl group. If the furthest chiral center has the same orientation as D-glyceraldehyde the sugar is designated as D; if the furthest chiral center has the same orientation as L-glyceraldehyde the sugar is designated as L. [1]



D-aldopentose L-ketohexose

Figure 2: Fischer projections with the furthest chiral centers highlighted

The Haworth projection is a simple representation of cyclic sugars in three dimensions which accurately displays the stereochemistry. The sugar ring is drawn planar and substituents are oriented perpendicular to the plane of the ring. Substituents can be displayed above or below the plane of the ring. However, Haworth projections are a poor representation of sugars as they do not provide any conformational information.

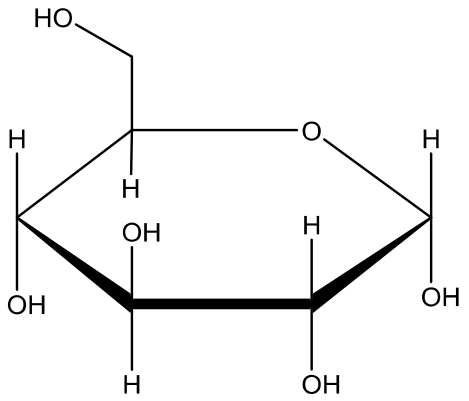


Figure 3: Haworth Projection of α-D-glucopyranose

Chair representations of sugars provide structural information about stereochemistry and conformation. Sugars are most commonly observed in the two common chair conformations; $^{4}C_{1}$ and $^{1}C_{4}$. In this notation, the letter C stands for a "chair" conformation while the numbers indicate the orientation (above or below) of the carbon atoms with respect to the reference plane of the chair which is composed of C_2, C_3, C_5 and the ring oxygen. D sugars are more stable in the $^{4}C_{1}$ conformation as this conformation minimizes the amount of bulky axial groups. [1]

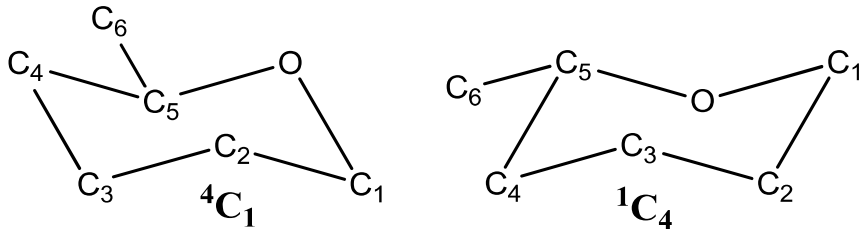


Figure 4: The two chair common conformations observed for the pyranose forms of sugars.

Monosaccharides exist in their chain or ring forms. The ring forms occur from the formation of intramolecular hemiacetals and hemiketals. These cyclization reactions occur spontaneously and reversibly in solution. For five and six carbon monosaccharides the ring forms are more common than the chain forms as shown in Figure 5. As an example, the most common form of glucose is a six membered ring. Rings are formed by the attack from a hydroxyl group on the fifth carbon of the aldose. The location of the attacking hydroxyl affects the formation of a furanose or pyranose ring. The aldehyde carbon is distinguished as the anomeric center. The nucleophilic attack on the anomeric center can occur from either face of the ring leading to formation of two possible diastereomers. These diastereomers are referred to as alpha and beta anomers. Alpha anomers have a cis configuration between the hydroxyl on C1 and the CH_2OH at C5 while beta anomers have a trans configuration between the hydroxyl on C1 and the CH_2OH at C5. [1]

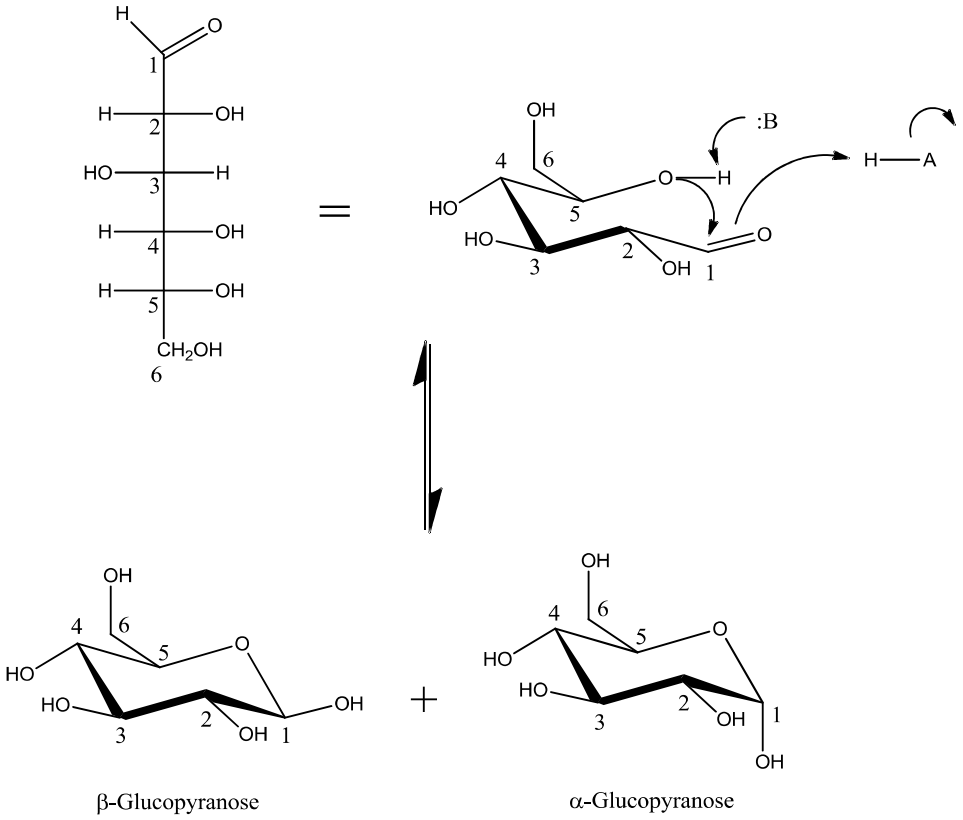


Figure 5: Hemiacetal formation

Monosaccharides undergo mutarotation in solution. Four conformations are possible as sugars can exist as either furanose or pyranose structures while the anomeric hydroxyl can be aligned axially or equatorially. Certain conformations will be favoured and more abundant for different sugars due to structural stability. Glucose is most commonly found as a pyranose structure that makes up more than 99 percent of the sugar in solution. The pyranose form is favoured as the bulky C6 group is placed in an equatorial position. β-Glucopyranose will be the most stable structure as the anomeric hydroxyl is placed in a more stable equatorial position. [1]

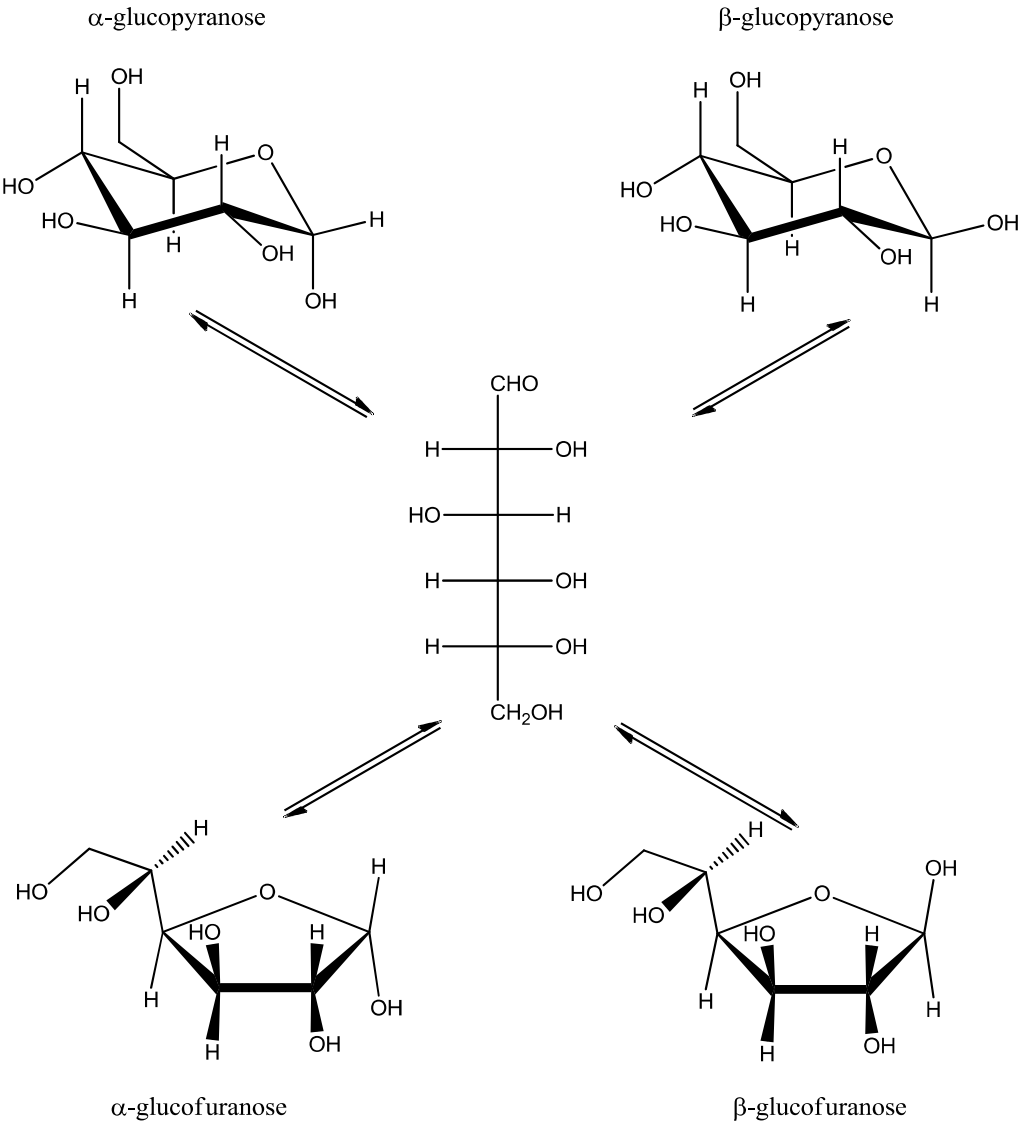


Figure 6: Mutarotation

Substituents on carbohydrates prefer to be in equatorial positions due to steric effects. Unintuitively, the anomeric substituent has been shown to prefer the axial position. The phenomenon is explained by the anomeric effect. The dipole moment from the endocyclic oxygen offsets the dipole from the hydroxyl in the axial position. If the anomeric hydroxyl is in an equatorial position the dipole moments are additive and destabilize the molecule. [2]

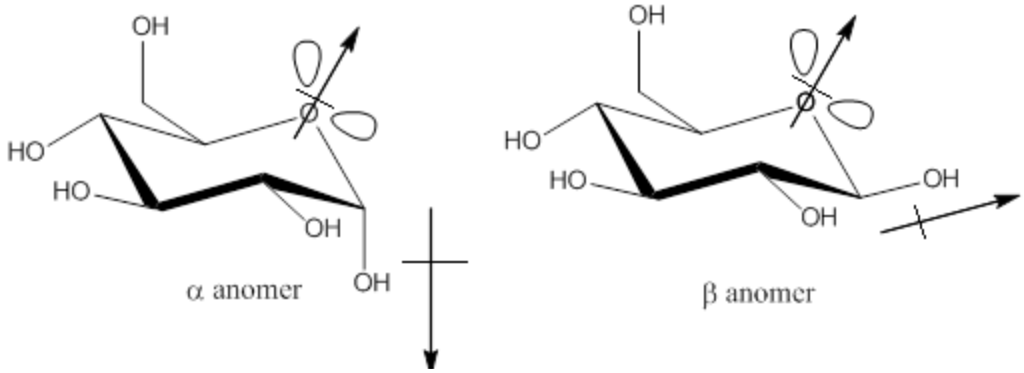


Figure 7: The Electrostatic anomeric effect

However, the electrostatic model runs into problems when dealing with the interaction between anomeric substituents and polar solvents. Increasing solvent polarity increases the presence of axial substituents. Using the electrostatic model, increasing the solvent polarity should increase the presence of more polar equatorial conformations. The molecular orbital model expands on the electrostatic model. In the molecular orbital model, the axial configuration is stabilized by the donation of electrons from an anti periplanar orbital on the ring oxygen. Stabilization of the equatorial configuration is not possible in this way as the oxygen orbital and anomeric carbon orbital must be anti-periplanar to one another. [2]

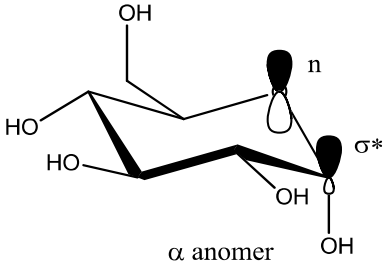


Figure 8: Molecular orbital anomeric effect

1.2 Carbohydrates in Nature

Carbohydrates are the most widely distributed, naturally occurring organic compounds on earth. [3] Photosynthesis produces carbohydrates as an energy source from light energy. Applications of carbohydrates are as structural materials in plants, animals and microorganisms. Proteins and lipids are linked to carbohydrates via glycosidic linkages. Glycolipids provide energy to organisms and can also serve as markers for cellular recognition while glycoproteins have a wide range of function including but not limited to immunological molecules, hormones and enzymes. [4]

Glucose (Glc) is the most common sugar found in living organisms. The abundance of glucose is attributed to the lower tendency of the residue to react non-specifically with amine groups on proteins. Glycation can impair or even destroy the function of many proteins. Lower rate of glycation can be attributed to the high stability of the cyclic form of glucose, meaning the molecule spends less time in the highly reactive chain form than other aldohexoses. [5] Glucose is an important energy source in humans and animals. Breakdown of carbohydrates in the body such as starch results in the formation of a lot of glucose. Through glycolysis, the citric acid cycle and oxidative phosphorylation, glucose molecules are oxidized to form CO_2 and water, yielding energy in the form of ATP. [6] Glucose is commonly found in linked forms such as: cellulose, starch, sucrose, and lactose.

Cellulose is the most abundant organic polymer on earth and the polymer has many significant biological and commercial applications. [7] The structure of cellulose is a polysaccharide consisting of a linear chain of several hundred to many thousands of β(1→4) linked D-glucose units.[8] Cellulose is an integral structural component of the primary cell walls

of green plants as it contributes to the rigidity of the cell wall. For commercial applications, cellulose is mainly used in the production of paper. [8]

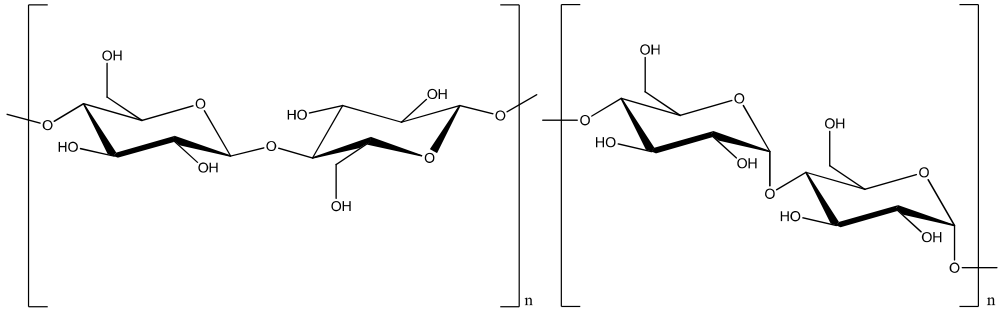


Figure 9: Cellulose repeating unit on the left. The repeating unit of Starch on the right.

Starch is similar to cellulose as it is comprised of (1→4) linked D-glucose units; however the glycosidic linkages are in the alpha conformation instead of the beta conformation. The conformational difference results in starch having different applications than cellulose. [9] Cellulose is responsible for structural stability while starch is applied for energy storage. Enzymes are capable of digesting starch, whereas cellulose cannot be digested as it does not fit the active site of these enzymes.

Polysaccharides can be found on the surface of bacterial cells. These polysaccharides are crucial for many cellular functions and aid the bacteria in survival from the surrounding environment.

1.3 Gram Negative and Positive Bacteria

Bacteria are distinguished by the composition of their cell well by using a technique known as Gram staining. Gram staining differentiates bacteria into Gram positive and Gram negative groups. Gram positive bacteria have a thick layer (50–90% of cell envelope) of peptidoglycan in their cell wall and are stained purple by crystal violet, whereas Gram negative bacteria have a

much thinner peptidoglycan layer (10% of cell envelope) meaning that they do not retain the purple stain and are counter-stained by Safranin to retain a pink color. [10]

1.4 Bacterial Cell Walls

Gram negative and Gram positive bacteria differ by the polysaccharide composition of their cell walls. Gram negative bactera contain lipopolysaccharides while sometimes containing capsular polysaccharides. The cell walls of Gram positive bacteria are composed of teichoic and lipoteichoic acids as well as sometimes containting a capsular polysaccharide.

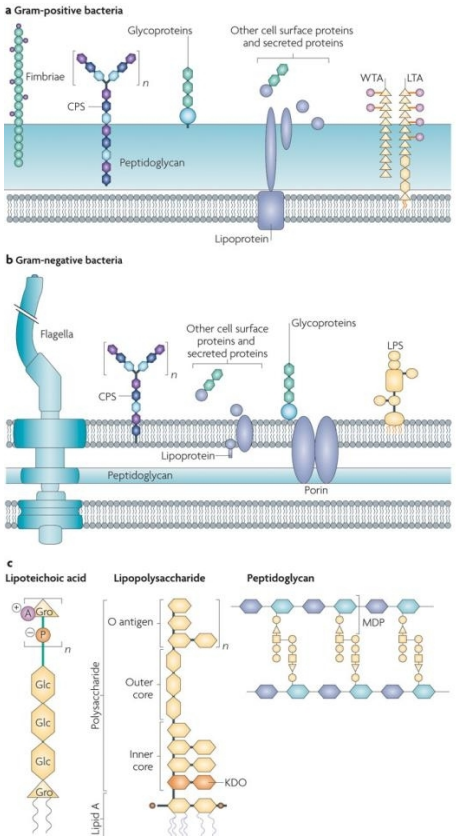


Figure 10: Gram Negative and Gram Positive Cell Walls. [11]

Adapted by permission from Macmillan Publishers Ltd: [Nature Reviews Microbiology] (Lebeer, S.; Vanderleyden, J.; De Keersmaecker, S. C. J. *Nat Rev Microbiol*, 2010, 8, 171-184.), copyright (2010)

1.41 Lipopolysaccharides

On the cell wall surface of gram negative bacteria there exist polysaccharides classified as lipopolysaccharides (LPSs). These molecules are composed of three different regions: the

hydrophobic Lipid A; the core oligosaccharide; and the O-antigen also known as the distal polysaccharide. [12]

The Lipid A unit is a fatty acid tail that is responsible for anchoring the LPS to the outer membrane of the bacteria. The Lipid A is an essential component of outer membrane structure and inhibition of Lipid A synthesis is lethal to Gram-negative bacteria. The structure is a glucosamine-based phospholipid structure that is highly conserved amongst all Gram-negative bacteria.

The structure of the lipid unit in Lipid A is the one that varies between bacteria. Characteristic features of the Lipid A structure in E coli are the two acyloxyacyl moieties on the lipid chain. This is shown in figure 11. [12] At the same time, the disaccharide unit that links the Lipid A to the core polysaccharide is a highly conserved structure. The disaccharide unit is typically a β-D-GlcN-(1→6)-α-D-GlcN disaccharide, which carries two phosphate groups. The Lipid A attachment point to the core oligosaccharide is this disaccharide unit.

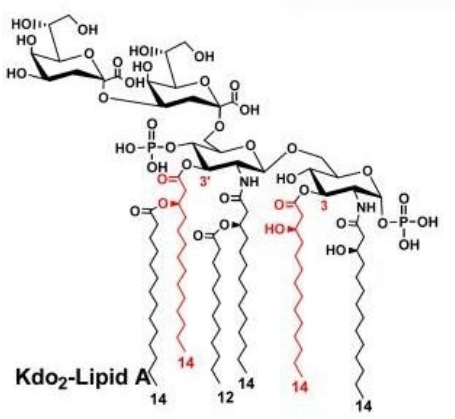


Figure 11: Lipid A structure linked to 2 Kdo residues from the core from *E. coli*. [12]

Adapted by permission from Annual Reviews: [Annual Review of Biochemistry] (Raetz, C. R. H.; Whitfield, C. *Annu Rev Biochem*, **2002**, 71, 635-700.), copyright (2002)

The core polysaccharide bridges the Lipid A anchor and the O-antigen repeating unit. Inner and outer cores make up the composition of the core polysaccharide. [13] The core polysaccharide is a non-repeating unit and similarities between these oligosaccharides are often

observed between members of the same genus. [11] The outer core consists of common sugar residues such as glucose, N-acetyl glucosamine, galactose, N-acetyl galactosamine; the overall structure of the outer core is more variable than the inner core. The inner core is made up of more "unusual" sugars such as various heptoses and ketodeoxyoctonate (Kdo). Kdo units are essential to the survival of the bacteria. [13] The Lipid A region is always linked to the core via Kdo units that are linked to the disaccharide glucosamine unit from the Lipid A. [13]

The final region of the LPS is the O-antigen. The O-antigen is placed furthest away from the outer membrane of the cell. The structure is a repeating unit, which can have subunits as large as fifty units. The units are very diverse, made up of a possibility of sixty different monosaccharides and thirty non-carbohydrate moieties. [11] O-Antigens differ from bacteria to bacteria and this variable repeating unit is responsible for the interaction between the cell and the host immune response; the structure is a contributing factor in the bacteria's ability to survive and evade an immune response of the host. [11] Identifying the structure of O-antigens is an important route in vaccine development against bacterial infections.

1.42 Teichoic and Lipoteichoic acids

Teichoic acids are found within the cell wall of gram positive bacteria and extend to the surface of the peptidoglycan layer. They are composed of glycerol phosphate or ribitol phosphate linked by phosphodiester linkages.[14] These bacterial polysaccharides are not found within the cell wall of gram negative bacteria. The main function of teichoic acids is to provide rigidity to the cell-wall. Teichoic acids are covalently linked to *N*-acetylmuramic acid of the peptidoglycan layer, to the lipids of the cytoplasmic membrane, or to a terminal D-alanine in the tetrapeptide crosslinkage between *N*-acetylmuramic acid units. Teichoic acids that are

anchored to lipids are referred to as lipoteichoic acids (LTAs), whereas teichoic acids that are covalently bound to peptidoglycan are wall teichoic acids (WTAs). [14] The difference between the two is displayed below in Figure 12.

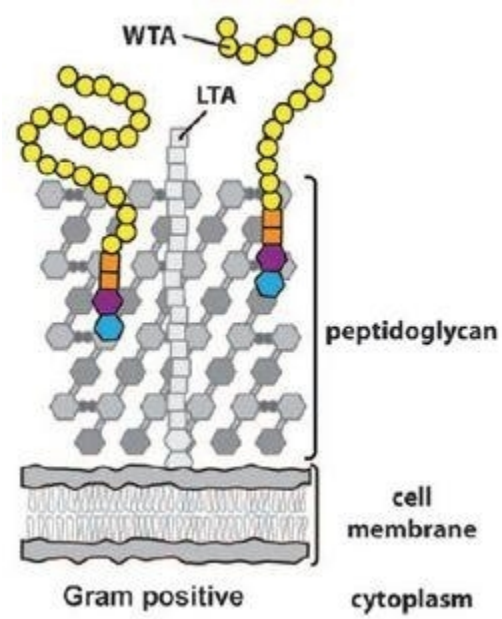


Figure 12: Composition of Gram positive cell wall showing LTAs and WTAs. [14] Adapted by permission from John Wiley & Sons, Inc.: [ChemBioChem] (Swoboda, J. G.; Campbell, J.; Meredith, T. C.; Walker, S. *Chembiochem*, **2010**, 11, 35-45.), copyright (2010)

1.4.3 Capsular Polysaccharides

Commonly, bacteria produce a layer of polysaccharide on the surface of the cell known as the capsular polysaccharide. Capsular polysaccharides provide bacteria protection from hostile environments and host immune defenses.[15] *C. difficile* has a polysaccharide capsule that may is involved in adhesion and evasion of the host immune system response through its anti-phagocytic properties. [16] The capsule protects the cell as antibodies produced by the immune

system do not bind efficiently to encapsulated organisms. Therefore, the capsule thus prevents recognition of the organism by the host immune system.

1.4.4 Lipooligosaccharides

The cell wall of *C. jejuni* does not produce any LPS. As a result, Lipooligosaccharides (LOSs) are produced in their place. LOSs are analogous to LPS. Their Lipid A structures are perform the same functions and conserve the same structure. However, the LOS lacks O-antigen units. [16] The structure of the core region is limited to ten monosaccharide units and is less conserved between LOSs than LPSs. Core oligosaccharides of LOS have been reported to be toxic and responsible for dangerous interactions with host immune systems. [17]

1.5 Importance of Characterization

Characterization of bacterial cell wall polysaccharides is crucial in the generation of vaccine candidates. It must be guaranteed that the polysaccharide and protein will be introduced into the human body safely without dangerous interactions with other common vaccines that are scheduled for infants and young adults.

In the case of *C. jejuni* , LOS structure is similar to oligosaccharides found on human gangliosides. If the LOS would be injected as a component of a vaccine; many complications would arise such as the autoimmune disorder: Guillan Barre syndrome. The immune response to LOS would produce antibodies that are not specific to *C. jejuni* LOS. These antibodies could attack host gangliosides that are crucial to motor function. However, the CPS is a safe component and must be purified from the LOS during extraction of the CPS. Characterization ensures that only CPS is used for the vaccine candidate. If the bacterial polysaccharide was not

characterized correctly, unwanted LOS could be incorporated into a vaccine candidate that would be dangerous to humans.

In the case of *C. difficile*, The surface polysaccharides that must be studied are the teichoic acids and the capsular polysaccharide.

Determining the structure of the polysaccharide for the vaccine candidate ensures the correct conjugation method is chosen. Knowledge of the location of specific functional groups is crucial in picking the correct conjugation method. The more that is understood about the structure the better the chance at successful conjugation.

1.6 *Campylobacter Jejuni*

C. jejuni is a Gram negative spiral shaped bacterium that forms commensal relationships in cattle, swine and birds. [18] However, the *campylobacter* species is the major cause of human bacterial gastroenteritis and is estimated to be responsible for 400-500 million cases per year worldwide. [19] Human illnesses are mostly associated with the *C. jejuni* and *C. coli* arising as food borne pathogens.

1.6.1 Campylobacteriosis

Campylobacteriosis is a collective description for infectious diseases caused by members of the bacterial genus *Campylobacter.* [20] Enteritis, caused by *C. jejuni* and *C. coli*, is the major public health concern caused by *Campylobacter* infection. Campylobacteriosis is primarily caused by the consumption of undercooked poultry, unsanitary drinking water and unpasteurized milk. [21] The infection rate of campylobacteriosis is significantly higher in developing countries than in developed countries. [20] Infection is a threat to tourists and soldiers, people

that are travelling to nations with high incidence rates of campylobacteriosis must be especially cautious of infection. [22] The disease is very prevalent worldwide as *C. jejuni* has a very low critical infective dose. Only 500-800 colony forming units (CFUs) are required to cause infection in humans while 10^{10} CFU are formed in the chicken intestinal tract for commensal relationship. [23] The infection can be treated with various antibiotics, erythromycin being the most common for children and tetracycline being the most common for adults. [24] Unfortunately, *C. jejuni* has been developing resistance against antibiotics; new methods must be used for dealing with campylobacter infection. [26]

Campylobacteriosis primarily leads to gastroenteritis and Travellers' diarrhea. The bacterium is the world's leading cause of gastroenteritis in developed and developing countries with approximately 400 million cases worldwide per year. [27] As well as, the infection being the second leading cause of Travellers' diarrhea in the world. [28] Stomach infections are usually self-limiting and the symptoms are cured after a few days. However, there are complications from *Campylobacter* infection that can result in paralysis. Complications arise as a result of autoimmune disorders caused by similarities between surface polysaccharides and gangliosides on the surfaces of nerve cells. Autoimmune disorders can start attacking the host's nervous system resulting in acute peripheral neuropathies such as Guillan-Barré syndrome (GBS) and Miller Fisher syndrome. Incidence of these neuropathies is quite rare, with GBS occurring at a rate of 30.4 per 100,000 people infected with *C. jejuni* . [29] *Campylobacter* infection is responsible for triggering about 25 percent of all cases of GBS. [29]

1.6.2 Surface Carbohydrates for *C. jejuni*

Bacteria are grouped together into serotypes based on the makeup of surface antigens. Relevant surface antigens for *C. jejuni* are the CPS and the LOS. However, LOS structures are similar to gangliosides, leading to severe complications with autoimmune disorders. Therefore, *C. jejuni* is typed by the composition of the CPS. Sugar chains of CPSs are unique to each subset of bacteria within a serotype. Serotypes can be classified by the passive haemagglutination test. [30] The test identifies serotypes of soluble heat-stable antigens by comparing antisera generated against each strain to the first known isolate of each serotype. [30] Antibodies are detected that were generated for fighting the relevant surface antigen. Consequently, the structures of the CPS within the same identified serotype have similar or identical structures. Several serotypes of *C. jejuni* have been characterized to date; the structures are shown in Table 1. The structures of the CPSs are very diverse displaying various pentoses, hexoses, N-acetylated hexoses and heptoses, all of which are commonly modified by O-methyl phosphoramidate. Existence of many serotypes emphasizes the need for a multivalent vaccine to fight *C. jejuni* . Prevention of infection hinders on understanding and being able to fight various CPS structures. Characterizing the structure of CPSs is the primary focus of the Monteiro research group as there exist many unique structures that need to be identified. Identification of the most common serotypes would lead to the development of a multivalent CPS conjugate vaccine that would be able to protect against all hypervirulent serotypes. [31] Production of a multivalent vaccine is seen as the most reasonable response to *C. jejuni* infection.

Penner Serotype	Serostrain	Structure	Discoverers
HS: 1	ATCC 43429	[-4)-α-D-Galp-(1-2)-(R)-Gro-(1-P]$_n$	Mcnally, D et al. (32)
HS: 2	NCTC 11168	6-OMe-D-glycero-α-L-glc-Hep-1 ↓ [2)-β-D-Rib*f*-1-5-β-D-Gal*f*NAc-1-4-α-D-GlcA6(NGro)- 1-]$_n$ 3 ↑ MeOPN	St. Michael et. al. (33)
HS: 3	BH0142	MeOPN ↓ 2 →4-α-Gal-(1→3)-L-glycero-α-ido-Hep-(1→ 3 ↑ Hydroxypropanol	Guerry, P. et al. (34)
HS: 4	CG8486	[→3)-6d-D-β-ido-Hep*p*-(1→4)-β-D-Glc*p*NAc-(1→]	Chen,Y et al. (35)

HS: 13	PG3019	MeOPN ↓ 2 or 7 [→4)-β-D-Glc*p*-(1→3)-6-d-α-D-ido-Hep*p*]$_n$	Unpublished
HS: 15	PG2887	[→3)-α-Ara*p*-(1→3)-6-d-α-gulo-Hep*p*(1→]$_n$	Unpublished
HS: 19	ATTC43446	[→4)-β-D-GlcA6NGro-(1→3)-β-D-GlcNAc-(1→]$_n$	Mcnally, D et al. (36)
HS: 23/36	81-176	[→3)β-D-GlcNAc-(1→3)-α-D-Gal-(1→2)-6d-α-D-altro-Hep-1→]$_n$ -6d-3OMe-α-D-altro-Hep-1→]$_n$ -D-glycero-α-D-altro-hep-1→]$_n$ -3OMe-D-glycero-α-D-altro-hep-1→]$_n$	Papp-Szabo E (37)
HS: 41	176-83	→2)-β-L-Ara*f*(1→2)-β-D-6d-altHep*f*(1→2)-β-L-6d-Alt*f*(1→ (75%) →2)-β-L-Ara*f*(1→2)-β-D-6d-altHep*f*(1→2)-α-D-Fuc*f*(1→ (25%)	Hanniffy, O. M. et al (38)
HS: 53	RM1221	α-D-Xlu 2 ↓ 2 P→3)β-6d-D-manno-Hep(1→3)α-6d-D-manno-Hep(1→3)α-6d-D-manno-Hep	Gilbert, M. et al. (39)

TABLE 1: Elucidated CPS structures isolated from various *C.jejuni* serotypes

1.7 *Clostridium difficile*

C. difficile is a Gram positive spore forming bacterium that is known to be a cause of enteric diseases in humans. [40] *C. difficile* can exist as spores which are metabolically inactive particles that are able to survive in soil, water, and on surfaces in clinical settings due to resistance against common sterilization methods such high temperatures, ultraviolet light and alcohol. *C. difficile* represents the leading cause of antibiotic-associated diarrhea and pseudomembranous colitis. In 2008, large outbreaks of *C. difficile*-associated diarrhea were reported internationally, and there have been reports of increases in severe disease, mortality and relapse rates. Currently, there is no commercial vaccine that affords protection against *C. difficile* and medical prevention is antibiotic based. However, antibiotics have led to an increase in the incidence of the disease. [41]

1.7.1 *Clostridium difficile* Infection

Spores represent the main vehicle for transmission, infection and persistence of *C. difficile*. If spores are ingested, they can survive in the stomach of infected subjects and eventually reach the intestine. The fate of these spores strongly depends on the environment provided by the host. Spores activate by a process termed germination. Germination can occur in response to many stimuli such as the bile acids present in the intestine. However, germination can also be prevented by other bile acids, such as chenodeoxycholate, which has a 10-fold higher affinity than the germinating bile acids for *C. difficile*.[40] The presence of chenodeoxycholate in association with aerobic conditions inhibits germination and growth of the bacterium during its passage through the small intestine. In absence or reduction of the normal commensal flora, spores become able to germinate into vegetative cells and the absence of competitors for

nutrients allows *C. difficile* to colonize in the colonic tract and cause infection. [40] Antibiotics that alter the composition of the commensal flora can be dangerous in allowing spores to germinate.

C. difficile is known to cause enteric diseases in many animal species, including humans. *C. difficile*-associated diarrhea (CDAD) is the most commonly diagnosed cause of hospital-associated and antimicrobial-associated diarrhea. [42] The risk of contracting CDAD is higher for elderly patients and those that have undergone hospitalization, gastrointestinal surgery, or those exposed to antibiotics. [43] In the United States, the estimated number of cases of *C. difficile*-associated disease exceeds 250,000 per year, with health care costs approaching US $1 billion annually. [44][45]

Since 2004, an unexpected increase in the incidence of CDAD has been observed. This has also been associated with higher rates of severe CDAD, treatment failure, and death. [46] Severe cases are being more frequently identified in younger patients and those patients without traditional risk factors. Many of these more severe cases have been linked to a new outbreak clone, designated ribotype 027. Prevention of *C. difficile* infection is based on patient isolation, improved sanitation, improved infection control, and antimicrobial restriction. Treatment of *C. difficile* infections is also becoming more problematic because the response to metronidazole, the main first-line treatment, is becoming unpredictable. The use of Vancomycin raises concern about the emergence of vancomycin-resistant enterococci and other vancomycin-resistant organisms. [47]

1.7.2 *C. difficile* Surface Carbohydrates

The chemical structures of two cell-surface polysaccharides produced by *C. difficile* are known and these cell surface polysaccharides are identified as PS-I and PS-II. [41] The structures are displayed in figures 13 and 14. *C. difficile* ribotype 027 or North American pulsotype 1 (NAP 1), which is the strain responsible for the widespread worldwide outbreak of *C. difficile*-related infection, was shown to express both surface polysaccharides: PS-I and PS-II. [18] *C. difficile* strain MOH900, also known as ribotype W or NAP2, is the most prevalent *C. difficile* strain in Ontario hospitals expresses the PS-II surface polysaccharide. Strain MOH718 which is an uncommon toxin variant also only expresses PS-11 on its surface. [41] PS-I is a pentasaccharide repeating unit containing three glucose residues and two rhamnose residues. PS-II is a hexasaccharide repeating unit with three glucose residues, two N-acetyl galactosamine residues and one mannose residue.

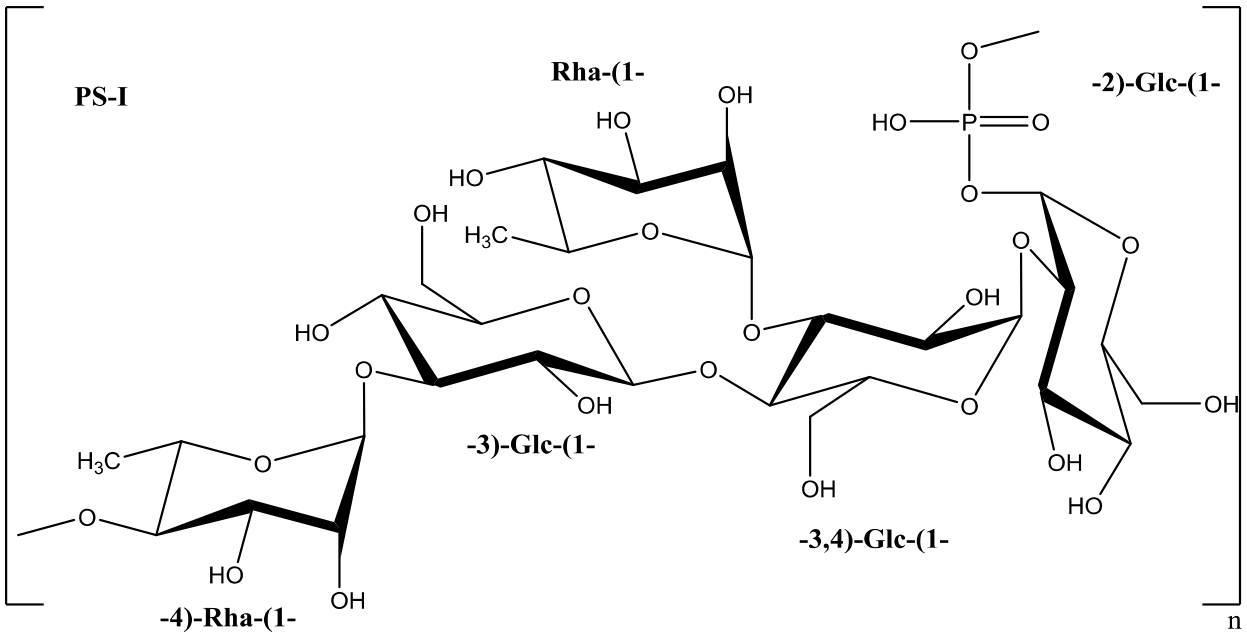


Figure 13: PS-I Surface Polysaccharide

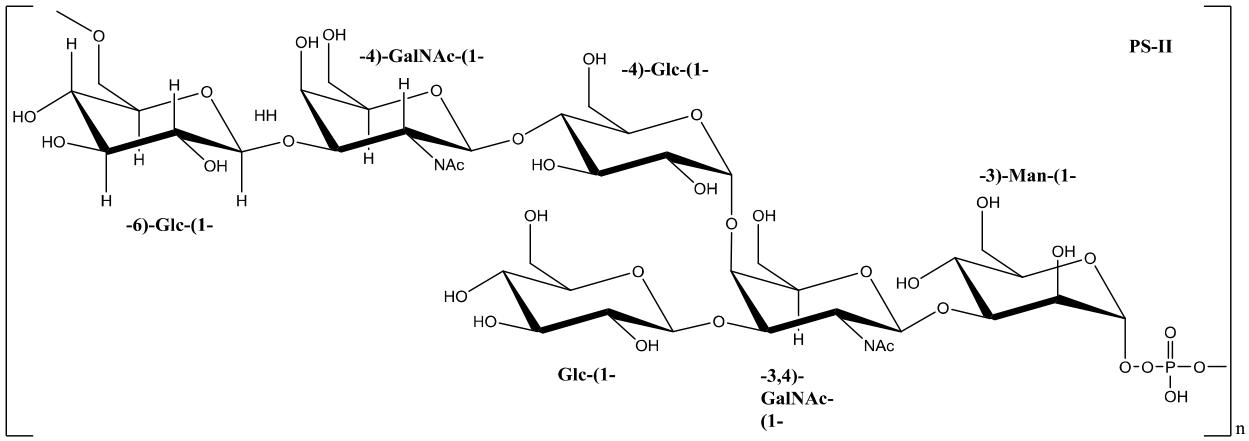


Figure 14: PS-II Surface Polysaccharide

1.8 Heptoses and Capsular Polysaccharides

C. jejuni is unique because it can synthesize heptose residues for its CPS. Serotypes containing: 6d-3OMe-altro-hep*p*, 6d-altro-hep*f*, 6d-ido-hep*p* and 6d-gal-hep*f* have been identified. The biosynthetic pathway of 6-deoxy-D-altro-heptose of serotype HS: 23/36, strain 81-176 has been discovered. [48] The researchers who looked at the pathway expected a simple pathway requiring sequential C6 dehydration, C3 epimerization, and C4 reduction of GDP manno-heptose all catalyzed by various enzymes. The pathway is shown below in Figure 15.

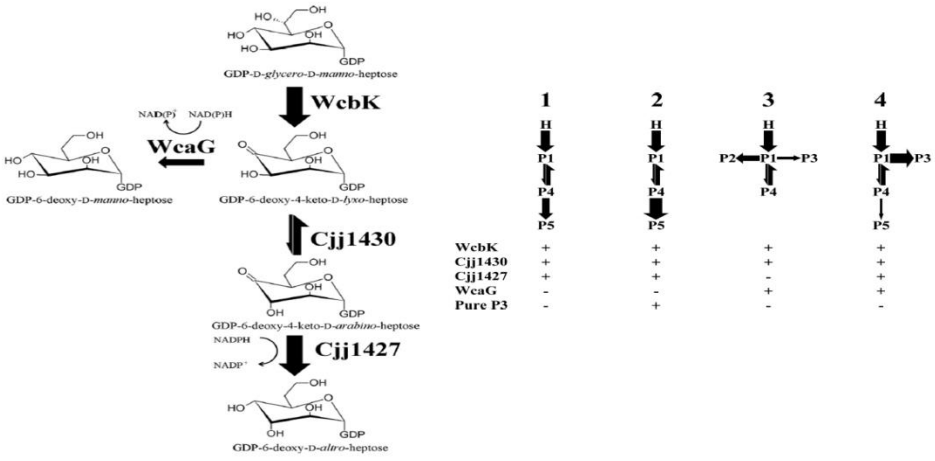


Figure 15: Synthetic Pathway for 6d-altro-heptose for Strain 81-176. Arrow thickness represents the efficiency of the enzymatic step. [48] This research was originally published in Journal of Biological Chemistry. McCallum, M.; Shaw, G. S.; Creuzenet, C. Comparison of Predicted Epimerases and Reductases of the Campylobacter jejuni d-altro- and l-gluco-Heptose Synthesis Pathways. *J Biol Chem*. **2013**; 288:19569-19580. © the American Society for Biochemistry and Molecular Biology.

Understanding this biosynthetic pathway is significant as it will allow for elucidation of similar heptose modification pathways; this is especially significant for *C. jejuni* because of the wide array of heptoses the bacteria can synthesize. Realizing the pathway also leads into a better understanding into the function of heptoses in CPSs. In the case of HS: 2 strain NCTC 11168 the capsule must biosynthesize GDP-3 6-OMethyl-L-gluco-heptose from the same precursor as HS: 23/36: GDP-D-glycero-D-manno-heptose. [49] The pathway is shown below in Figure 16.

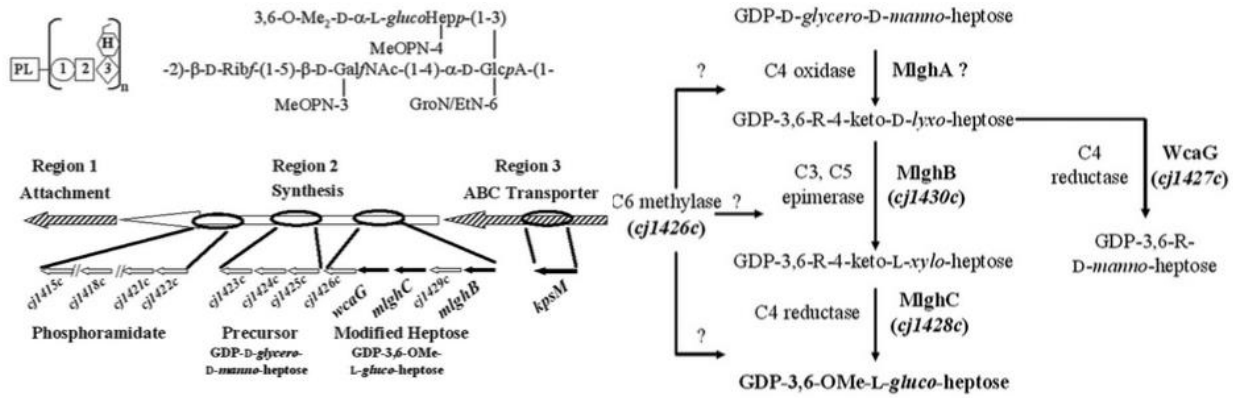


Figure 16: Synthetic Pathway for GDP-3,6-OMe-L-gluco-heptose for Strain NCTC 11168. [49] Adapted by permission from John Wiley & Sons, Inc.: [Mol. Microbiol] (Wong, A.; Lange, D.; Houle, S.; Arbatsky, N. P.; Valvano, M. A.; Knirel, Y. A.; Dozois, C. M.; Creuzenet, C. *Mol. Microbiol*, **2015**, 96, 1136-1158.), copyright (2015)

It is evident from the pathway that enzymes cj1427c, cj1430c and cj1428c, which are encoded by the genes wcaG, MlghB and MlghC, are responsible for synthesizing modified heptoses for the *C. jejuni* CPSs. (41) Mutants of the strain that did not contain these genes were studied for the effects of not being able to synthesize heptoses for the CPS. It was found that heptose modifications render strains more susceptible to bile salts than strains that were not capable of synthesizing the heptose modifications. The mutants, to make up for the lack of heptoses, modify the structure of the rest of the CPS. Variations in the amount of MeOPN, GroN, and EtN alter the hydrophobicity of the surface of the capsule to increase protection against bile

salts. The CPS protects the cell against serum killing in the host. Serum complement is an essential innate immune defense against pathogens and the sensitivity of a strain to the serum is determined by the composition of the CPS. [49] Heptose modification mutants showed that modified heptose is essential for protecting bacteria against serum as all the modified mutants suffered more susceptibility to the serum than the wild type. Modifying the heptose and MeOPN levels did not affect the bacteria's interaction with macrophages. However, the tests were performed on murine macrophages and different effects could be possibly observed when dealing with human or avian macrophages. This may explain how *C. jejuni* is pathogenic is certain organisms while commensal in others. Interfering with the heptose biosynthesis affects the ability of the cell to invade the host. The wcaG::catΔ mutant experienced increased invasion despite lacking heptoses and MeOPN. However factors may vary from strain to strain as in strain 81-176 the MeOPN contributes to invasion of epithelial cells. The variations of invasion correlated well with the levels of transcription of cj1429c in the mutants and may be due to additional surface differences mediated by cj1429c. [49] Understanding the heptose modification pathway is important for understanding host colonization and pathogenicity.

The role of the CPS and the heptose modification pathway on chicken intestinal tract was investigated. Mutants lacking the wcaG gene were impaired in colonization (number of chicks colonized) and persistence (severity of infection) when compared to the wild type. The mlghB mutant showed problems with persistence of infection. [49] Overall fine tuning of the CPS structure, using heptose residues and MeOPN, allow for the cell to modify its structure to reach an optimal configuration in surviving against the various host defenses found in the gastrointestinal tract. (bile salts, serum and phagocytic cells) The capsule structure is crucial in contributing in the strength and persistence of colonization of the intestinal cells. Studying the

various heptoses that can be biosynthesized for the capsule and how they affect the cells can only further understanding the pathogenicity of *C. jejuni* which can only aid in vaccine development.

1.9 Vaccinology

Edward Jenner developed the first vaccine to combat smallpox in 1796. Jenner realized that milkmaids that had been infected with cowpox were immune to smallpox outbreaks, leading Jenner to the discovery that inoculating patients with cowpox led to immunity from smallpox. [50] Hence, the term vaccine is derived from the name of smallpox of the cow: *Variolae vaccinae*. It was not until 1885 that Louis Pasteur had developed a second vaccine, this one aimed at rabies infection. Pasteur produced a rabies antitoxin that functioned as a post-infection antidote. Pasteur expanded the term vaccine past cowpox and smallpox to all inoculating agents. [52] Today's definition of a vaccine is "a suspension of live (usually attenuated) or inactivated microorganisms (e.g., bacteria or viruses) or fractions thereof administered to induce immunity and prevent infections disease or its sequelae." [53]

To simplify, vaccines use a weakened pathogen to illicit the immune response into producing protective antibodies that protect against subsequent infection. If a host's immune system has been primed by vaccination, the immune system recognizes the pathogen, eliminates it and is able to recognize the pathogen in the case of re-exposure. Re-exposure to the pathogen leads to a significantly shorter response time from the immune system. The approach is known as active immunity and differs from passive immunity in that the body produces its own antibodies instead of being given antibodies. Passive immunity is not viable as it lasts only a few weeks or months while active immunity is a long lasting solution. Diseases that have been prevented using

vaccines are: smallpox, diphtheria, measles, mumps, pertussis, paralytic polio, rubella, tetanus, and invasive hemophilus influenza. [54]

Vaccines must be produced with extreme care as some pathogens are dangerous to the host if similar structures are included in the vaccine. Bacteria have lipid A components which are responsible for linking the surface polysaccharides to the membrane of the bacteria. Lipid A is enterotoxigenic in humans, rendering it unsafe as a component of vaccines. *C. jejuni* contains lipooligosaccharides that are similar to human gangliosides that can lead to autoimmune complications. The immune system attacking its own host due to similarity between pathogen/host cells is known as an autoimmune disease. Some pathogens such as HIV are too dangerous to be used as antigens even as attenuated viruses due to the high mutation rate and genetic variability.

1.9.1 The Immune System

The immune system acts using two mechanisms: innate and adaptive immunity. Innate immunity is the first contact point when the body is attacked by a pathogen. The response is not pathogen specific and the body uses a set of pre-existing protective molecules to fend off general pathogens [53]. The innate immune system, as shown in Figure 17, comprises anatomical barriers against infection, both physical and chemical, as well as cellular responses. The main physical barriers are the skin and mucosal membranes of the body. Epithelial barriers prevent infection by simply not allowing pathogens to enter the body. [53] There are chemical barriers on these surfaces that are specialized substances that possess antimicrobial activity as well as an acidic pH. Pathogens that breach physical and chemical barriers, due to damage or direct infection, can survive in extracellular spaces or even infect other cells of the host, eventually

replicating and spreading to other parts of the body. [53] The cellular innate immune response responds to invasions that overcome epithelial protection. Response is rapid, usually beginning minutes after invasion. The response is triggered by cell surface or intracellular receptors that detect pathogens based off specific molecular components. White blood cells such as macrophages and neutrophils are activated and engulf extracellular microbes through the process of phagocytosis. [53] The body starts producing proteins and other helpful molecules that have antibacterial activity or the ability to recruit fluid, cells and molecules to sites of infection. The result is inflammation as the influx causes swelling and other physiological changes. Local innate and inflammatory responses are beneficial for eliminating pathogens and promoting healing for damaged cells while removing dead cells. For example, natural killer cells are lymphocytes that are responsible for recognizing and killing virus-infected, altered or stressed cells. Phagocytic cells are responsible for eliminating the pathogens and dendritic cells take up pathogens and present them to lymphocyte to activate the adaptive immune response.

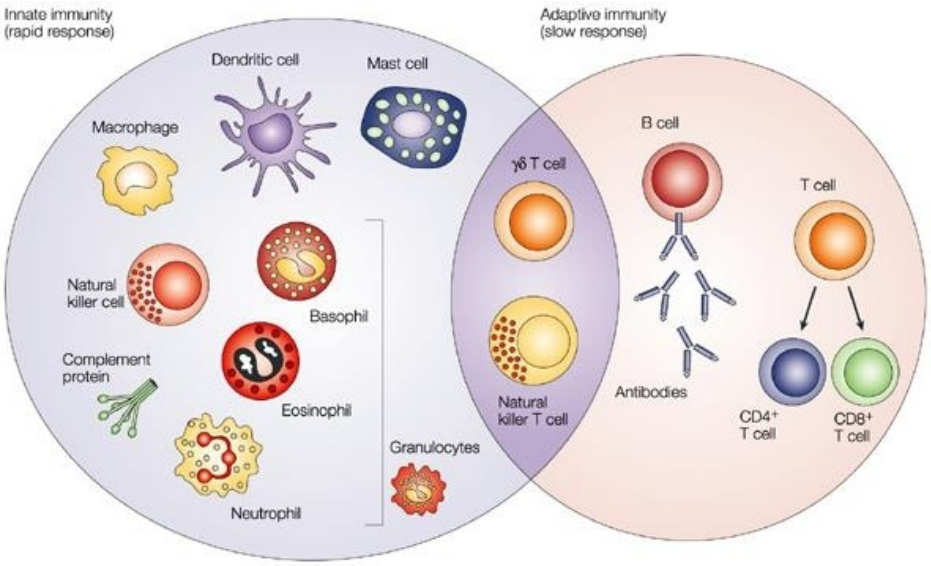


Figure 17: The innate immune response [53]

Adapted by permission from Macmillan Publishers Ltd: [Nature Reviews Cancer] (Dranoff, G. *Nat Rev Cancer*, **2004**, *4*, 11-22.), copyright (2004)

Adaptive immunity is specific to pathogens the immune system is trying to defend against, so the immune system must be capable of identifying the specific pathogen for selective elimination. The adaptive immune system must be able to identify a diverse set of specific antigens with extreme precision, must produce immunological memory from the presented antigens and must be able to distinguish between host and non-host cells. The immune system detects very subtle differences between antigens. For proteins; the difference between amino acid sequences can be pinned down to one a one amino acid difference. [53] The innate immune system does not have such strong recognition capability. The adaptive immune system establishes immunologic memory for specific antigens. Memory is established for secondary attacks and makes fighting secondary attacks quicker and easier. [53] Memory response is long term and can remember many infectious pathogens. Recognizing antigens that belong to the host and non-host is very important because surface pathogens can be very similar to molecules on the surface of host cells. If the immune system was not able to distinguish between host/non-host, responses to host components could lead to serious complications or even death. [53]

To generate an effective immune response, the immune system enlists the help of two important cells: antigen presenting cells (APCs) and lymphocytes. There are 3 types of lymphocytes: natural killer cells, B lymphocytes (B cells) and T lymphocytes (T cells). Natural killer cells are important for innate immunity while B and T cells are integral in adaptive immunity. B cells are developed in the stem cells of the bone marrow. Millions of different B cells are produced each day. B cells are capable of producing immunoglobulins that are specific to chemical structures, and effectively all antigens that belong to infectious microorganisms. However, B cells are not capable of producing antibodies until they are fully activated. Each B cell has a unique receptor protein that binds to one specific antigen in order to activate the B cell.

The activated B cell undergoes rapid division and creates memory B cells and plasma B cells. Division is done because memory B cells are responsible for displaying antibodies on the surface, like the parent B cell, while plasma B cells are responsible for secreting antibodies. Memory B cells survive for a long period of time while the plasma cells have a shorter lifespan of a only a few days in which they secrete hundreds to thousands of antibody molecules. [28]

Like B cells, T cells also grown in the bone marrow. However, T cells mature in the in the thymus gland which B cells do not. [53] After maturation there exist two well defined types of T cells; the T helper (T_H) and T cytoxic (T_C) cells. The two subpopulations are differentiated by the membrane glycoprotein found on the surface of the cells. T helper cells are known as CD4+ and the T cytoxic cells are known as the CD8+ cells. While B cells are able to recognize free antigens, T-cell receptors can only recognizes antigens that are positioned on the surface of other cells. Antigens must be held within the binding groove of a cell surface protein known as the major histocompatibility complex (MHC) molecule. There are two classes of the MHC molecule: class I and class II. Class I molecules specialize in presenting antigens that originate from the cytosol. The antigens are presented to the $CD8^+$ T cells which recognize and kill cells expressing the antigen. Class II MHC molecules are found on leukocytes called antigen-presenting cells (APCs). Class II specializes in presenting antigens from bacteria that have been engulfed by various cells. B cells express the class II MHC molecules. Once the antigen is expressed on the surface it is presented to $CD4^+$ T cells which become activated and produce immunity against extracellular invaders.

1.9.2 Polysaccharide Vaccines

Initially, the idea that polysaccharides could be used as vaccine candidates was founded on the basis that pneumococcal polysaccharides were found to be immunogenic. [54] The vaccine was the first to be derived from a capsular polysaccharide. A tetravalent vaccine was first used in 1945. [54] In the 1970's the vaccine was improved to a 14-valent vaccine and in 1983 this evolved into a 23-valent vaccine. [55] This is the current vaccine that is used for pneumococcal disease and it is known as Pneumovax 23 (PPV-23). However, this vaccine utilizes a polysaccharide antigen without a conjugated carrier protein. Not utilizing a carrier protein leads to problems with the immunity of the vaccine. Polysaccharide antigens are not processed by antigen processing cells (APC) but interact directly with B cells, inducing antibody synthesis in the absence of T cells. [56] T cell independent responses are limited in a number of ways. Most importantly, the response in young children below 18 months of age is not significant and does not induce significant or sustained amounts of the required antibodies. There are also problems with geriatric patients. [56] Polysaccharide vaccines are immunogenic in older children and adults however the characteristics of the immune response are limited and often short lived. The antibodies produced by B cells are: IgM and IgG2, which do not have a long life time and have trouble being elicited on repeated exposure to the pathogen. The inability of polysaccharide antigens to induce immunological memory is a problem that is solved by utilizing protein antigens. Antibody responses to protein antigens must utilize T cells. T cell dependant responses are elicited in very young children and the immunity is long lived due to the production of the IgG1 and IgG3 antibodies. Immunity is long lived and immunological memory is formed. [56]

1.9.3 Conjugate vaccines

Conjugate vaccines are created by covalently attaching a poor antigen, usually a polysaccharide, to a carrier protein, thereby improving the immunological attributes of the poor antigen. This process is shown in figure 18. When the protein is conjugated to the polysaccharide and used as a vaccine, CD4+ T cells are generated as a part of the immune response. [57] T cell responses allow for the generation of high affinity antibodies that will ultimately lead to long-lived antibody mediated protection. Commonly used proteins for conjugation are: tetanus toxoid (TT), diphtheria toxoid (DT), mutant diphtheria toxoid (CRM_{197}), and outer membrane proteins. [56] Carrier proteins produce an antigenic response dominated by IgG1 and IgG3 antibodies. IgG1 and IgG3 bind with high affinity to phagocytic cells priming the immune system for memory response. [56] The interaction was first noted by Avery and Goebel, in 1929, and was the landmark for the development of conjugate vaccines [56].

The first commercially available glyconjugate vaccine formulated for use in humans was for use against *Haemophilius influenza* type B (Hib). [56] The vaccine was introduced into the infant vaccination schedule and was successful in reducing the Hib disease in children. After routine use of the Hib vaccine in the U.S. from 1980 to 1990, the incidence of invasive Hib disease was decreased from 100 infections per 100,000 children down to 1 per 100,000. [56] The success of this vaccine spearheaded interest in developing other vaccines for other encapsulated bacteria that are the cause of various childhood diseases. Other glyconjugate vaccines that have shown success commercially are for *Steptococcus pneumonia* (23 serotypes), and *Salmonella typhi*.

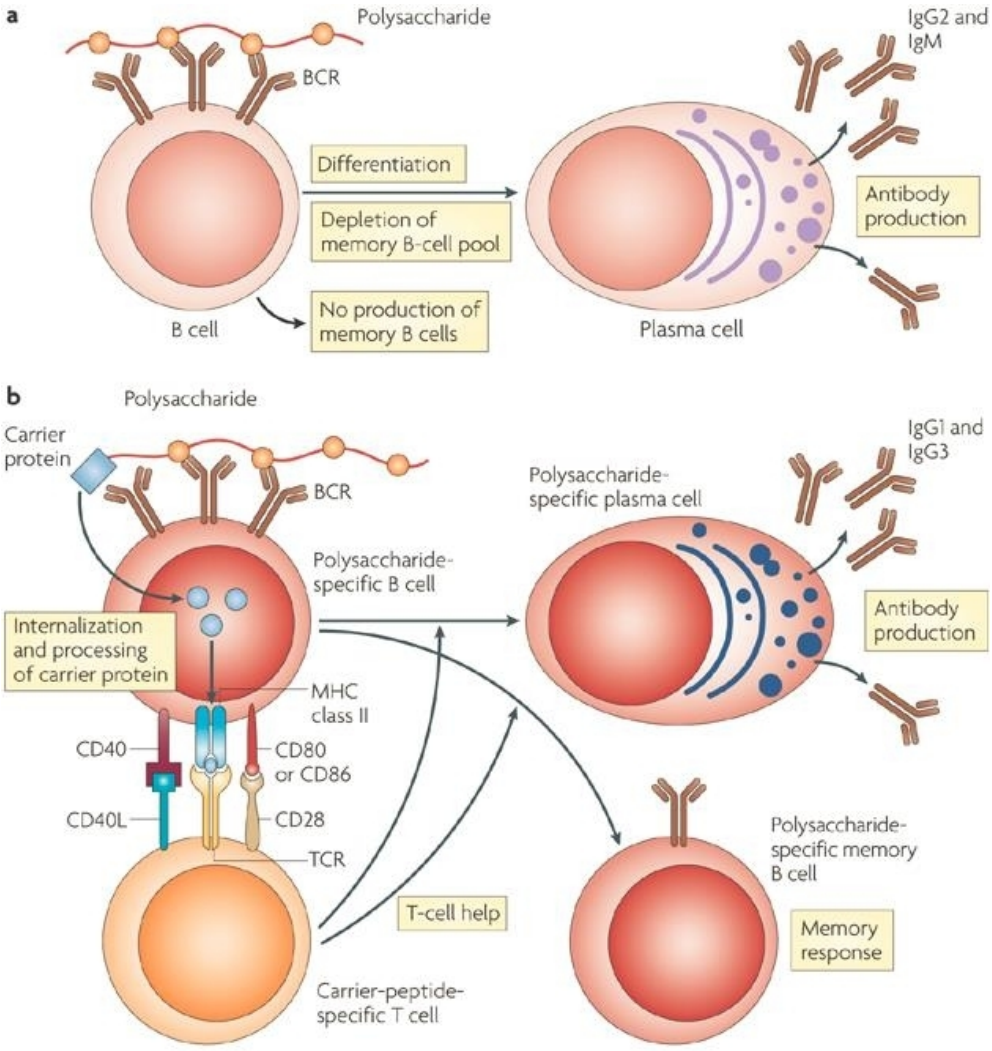


Figure 18: The immune response for a pure polysaccharide compared to the response for a polysaccharide conjugated to a protein. [58]

Adapted by permission from Macmillan Publishers Ltd: [Nature Reviews Immunology] (Pollard, A. J.; Perrett, K. P.; Beverley, P. C. *Nat Rev Immunol*, **2009**, 9, 212-220.), copyright (2009)

1.9.4 Ongoing development on a multivalent CPS glycoconjugate anti-*Campylobacter jejuni* vaccine.

The target population for a conjugate vaccine against *C. jejuni* would be civilian and military travellers to endemic regions and pedriatic populations living in endemic regions. [59] Extensive data is available on the most popular Penner strains of *C. jejuni* from the developed world, however there is much more limited data from less developed countries (LDC) where the

incidence of *C. jejuni* is high. Figure 19 summarizes data on the Penner serotypes from sporadic clinical cases based on >16,000 clinical isolates from the developed world and about 700 from LDC. [59] HS: 1 HS: 2 and HS: 4 are the most common serotypes encountered globally. HS: 3 HS: 5 and HS: 8 are also encountered on the global scale, while other serotypes tend to show geographic and population based variability. [59] More than ten percent of the serotypes from the developed world and more than 20 percent from the developing world could not be typed using Penner serotyping. This could be because of CPS not covered by Penner serotyping or that Penner serotyping requires a CPS for typing and some cells can turn off capsule production by phase variation. [59]

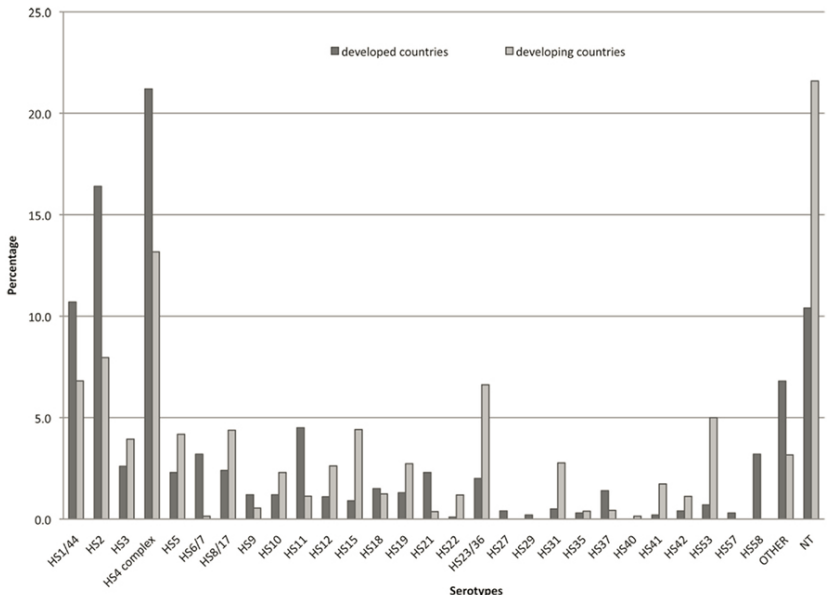


Figure 19: Penner serotyping for *Campylobacter jejuni* in developed and non-developed countries. [59]

serotypes HS: 3, HS: 4, HS: 10, HS: 13 and HS: 23/36. The overall goal is to combine all the glycoconjugates into one multivalent vaccine which would attempt to provide immunity against all common serotypes of *C. jejuni* . The HS: 23/36 glycoconjugate has been tested in the New World monkey: *Aotus nancymaae* which develops diarrheal disease that is similar to the human

illness following *C. jejuni* infection. [59] Immunization with CPS–CRM197 was attempted to see if the vaccine could prevent diarrheal disease, monkeys were immunized with 2.5 μg of polysaccharide adjuvanted with alum three times with an interval of 6 weeks between subcutaneous injections. Nine weeks after the third immunization, monkeys were challenged with approximately 1011 CFU of *C. jejuni* 81–176. The diarrheal attack rate for non-immunized control monkeys was 70%. However, 100% of vaccinated animals (14/14) were protected against diarrheal disease. [60] The work was able to demonstrate for the first time that a conjugate vaccine using a CPS was capable of protecting against enteric disease and proved that the capsule provides a crucial role in the bacteria's virulence.

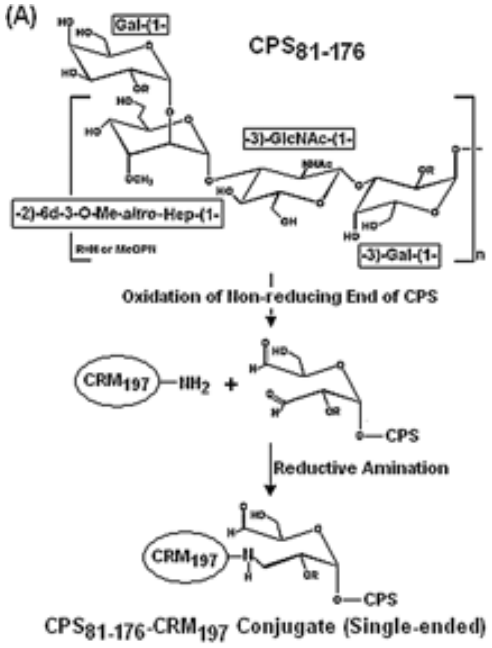


Figure 20: Scheme for synthesis of the HS: 23/36 conjugate vaccine. [59]

1.9.5 Development of an anti-*C. difficile* glycoconjugate vaccine.

PS-II was found to be the most abundantly expressed surface polysaccharide by the most common ribotypes of *C. difficile* including the hypervirulent (NAP1/027) and other clinical isolates belonging to ribotypes 001, 018, 027, 078 and 126. [40] However, the polysaccharide coating was found to be not as thick and uniformly distributed as expected for a capsular polysaccharide compared to other CPSs analyzed in the past. Therefore, it can by hypothesized that PS-II is expressed not as a capsule but as a cell wall linked polysaccharide not bound to the peptidoglycan or as a conjugate bound to the lipoteichoic acids. Antibodies against the conjugated PS-II were able to detect the polysaccharide on the bacterial vegetative cells, thus confirming that PS-II is a suitable target molecule for vaccine research. [61] More encouragingly, a recent pre-clinical protection study showed that a conjugate vaccine, composed of PS-II adjoined to the immunostimulatory protein keyhole limpet hemocyanin (KLH), protected 90% of mice when challenged with an LD50 dose of *C. difficile* spores. [62]

CHAPTER 2: Scope of Research

2.1 Information Needed for the Characterization of Surface Polysaccharides

Many pieces of information are required to fully identify the composition of a bacterial capsular polysaccharide. The composition of monosaccharides and how they are linked together must be determined. The enantiomeric configuration must be established, as well as the anomeric configuration of each monosaccharide and its linkage. The last step is to discover the presence and location of all non-sugar functional groups, if they exist.

2.2 Aim of the Study

This research involves the isolation, purification and characterization of the capsular polysaccharide in *C. jejuni* serotype HS: 44 strain PG 2871. This strain was chosen because of its interesting structure that comprises of two capsular polysaccharides. One capsule is thought to have the same CPS structure of HS: 1 strain PG 856 while the second capsule is a new capsule structure made up of 6-deoxy-heptoses. The goal was to separate the two capsules using a base treatment and characterize the structure of the second capsule. The strategy of characterization will employ a number of analytical and chemical techniques such as: derivatization, gas chromatography – mass spectrometry (GC-MS) and nuclear magnetic resonance (NMR) experiments. Upon successful characterization, the objective would be to conjugate the CPS to an appropriate protein carrier for use in a multivalent anti-*Campylobacter jejuni* glycoconjugate vaccine.

Here, we also aimed for the detection of mannan units in *C. difficile*, discovered when *C. difficile* biomass was exposed to extended hot water-phenol treatment. Mannans are linear polymers of mannose. The analysis revealed a preparation that comprised a high mannan

component. We also set out to determine if there were differences in the amount of mannans when analyzing preparations with high spore content.

During prolonged hot water-phenol extraction, the Man-rich material co-solubilized with the previously described water-soluble polysaccharides and in most cases it proved difficult to isolate it by size-exclusion chromatography. Nevertheless, enough quantities were attained that allowed for trustworthy analysis by gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-MS (LC-MS).

CHAPTER 3: Materials and Methods

3.1 Bacterial Growth

Capsular polysaccharides are isolated from inactivated bacteria and then must be purified from unwanted portions of the bacterial cell wall such as Lipid A and LOS. The Monteiro research group received the *C. jejuni* HS44 PG 2871 sample from the Enteric Diseases Department of the US Naval Medical Research Center. The bacterial sample was grown in brain heart infusion medium at 37°C in a microaerophilic environment and inactivated by sonication to produce approximately 25 grams of wet cell pellet.

A fresh toxigenic *C. difficile* strain isolated from a horse with clinical signs of gastrointestinal disease was grown anaerobically for 37°C for 5 days in dialysis tubing suspended in 4L flasks of brain heart infusion broth. An overnight culture (1 mL) of this toxigenic strain was used to inoculate a sterilized dialysis bag (MWCO 10,000 Da) containing 500 ml PBS before sterilization, suspended in 3- 3.5 L of brain heart infusion broth, in a stoppered, vented 4L conical flask. After 5 days of incubation at 37°C in an aerobic incubator, the contents of the dialysis bag were harvested (10,000g, 30 min) and the resulting pellet was stored at 4°C. Five-day cultures were found to be richer in spore content (compared to 3-day-old cultures). The pellets were pooled for carbohydrate extraction.

3.2 Extraction and Purification

Once the intact cell mass is obtained, purification of the CPS can begin. The first step is a water/phenol extraction that separates the bacterial surface carbohydrates, which are found in the water layer, from lipids and proteins, which are separated into the phenol layer. The procedure involves crushing the sample and adding it to a 500 mL mixture of a 40% phenol: 60% deionized

water solution. The mixture is heated to 75° C for 6 hours with constant mixing. Upon cooling in an ice water bath overnight, the layers separate into the water layer being on top leaving the phenol layer on the bottom. The water layer is collected, and replaced with a fresh volume of water, and the procedure is repeated two more times to increase the yield of polysaccharide obtained. [63]

When the water layer is obtained it is still saturated with phenol from the extraction that must be removed. Phenol and other small molecules are removed by placing the sample in a Spectrapor 1kDa molecular weight cut off (MWCO) dialysis membrane in a beaker of running deionized water for 24 hours. The phenol free dialyzed extract is frozen, lyophilized and dissolved in a small amount of water for ultracentrifugation at 44,000 RPM for 24 hours. The ultracentrifugation is performed to separate a CPS containing supernatant from a pellet that consists of unwanted core polysaccharide. However, the ultracentrifugation does not remove all the unwanted core polysaccharide. The supernatant is then purified by size exclusion chromatography on a Biogel P-2 hydrophilic column with a nominal exclusion limit of 100-1800 Da to separate the remaining core polysaccharide from the CPS. The column produces small 1 mL fractions which are frozen and the lyophilized to isolate the polysaccharides. The CPS is expected to elute before the core polysaccharides as it is larger and cannot penetrate the pores of the beads of the column as easily. Therefore, the CPS is known to elute immediately after the void volume.

The *C. difficile* biomass was stirred for 5h at 71 ^{o}C in a water-phenol solution (60%:40%). After cooling the preparation parted into aqueous and phenol layers. The water layer was dialyzed (MWCO 1000), lyophilized and applied to a Sephadex-G50 column (1 cm x 1 m), in which a limited quantity of the mannan eluted after polysaccharide PS-II.

3.3 Gas Chromatography Mass Spectrometry and Liquid Chromatography Mass Spectrometry

Gas chromatography is a separation technique where the components of a vaporized sample are separated between a mobile gaseous phase and a liquid stationary phase held in a capillary column. [64] The samples analyzed by gas chromatography must be volatile and thermally stable. The mobile phase in gas chromatography is referred to as the carrier gas and it must be chemically inert. Helium is the most common, although argon, nitrogen and hydrogen are also used. To ensure efficient injection the sample is introduced to the column as a "plug" of vapour. Slow injection or an oversized sample can result in spreading of bands and/or poor resolution. Liquid samples are injected through a septum into a heated sample port. The sample port is kept 50°C higher than the boiling point of the least volatile component of the sample. [64] Capillary columns are the most common columns and they are made of fused silica. They only require a small amount of sample compared to old packed columns and therefore require a sample splitter to deliver a small split amount of sample (1:100 to 1:500), with the remainder going to waste. The temperature of the column is a very important variable that must be controlled very precisely resulting in columns being housed in thermostated ovens. Temperature programming allows for the temperature of the column to be controlled during separations where the temperature is increased continuously or in steps as the separation proceeds. Optimal resolution is usually attained with minimal temperature however the lower the temperature the longer the separation time. Temperature programming allows separations to have good resolution for key regions of separation while lowering the overall time for separation. [64]

The flame ionization detector (FID) is most widely applied detector for gas chromatography. With a FID, the effluent from the column is directed into a small flame made

from air and hydrogen. Most organic compounds are pyrolyzed by this flame resulting in the production of ions and electrons. Compounds are detected by monitoring the current produced when the ions and electrons are collected. The number of ions produced is roughly proportional to the number of carbon atoms in the flame. [64] The detector is mass sensitive and not concentration sensitive as it responds to the number of carbon atoms that enter the detector per unit time. The FID is high sensitive detector with a large linear response range and low noise. However, it is destructive to the sample and requires additional gas controllers to provide the air/hydrogen mixture for the flame.

Gas chromatography can be coupled with a mass spectrometer to produce a very powerful analytical technique known as GC-MS. [64] Mass spectrometers detect the mass to charge ratio of ions that have been produced from the analyte. Capillary columns, due to their low flow rates, allow for the column output to be fed directly into the ionization chamber of the mass spectrometer. The most common ionization techniques for GC-MS are electron impact and chemical ionization and the most common mass analyzers are quadropoles or ion traps. GC-MS will produce a mass spectrum for every compounds separated via chromatography. Compounds are identified based on the mass spectra and not solely on the retention time of the components. Applications of GC-MS include drug detection, fire investigation, environmental analysis, explosives investigation, and identification of unknown samples.

Liquid chromatography mass spectrometry is a technique that couples the physical separation capability of liquid chromatography or high performance liquid chromatography with the mass analysis capability of mass spectrometry. Liquid chromatography is a separation technique in which the utilized mobile phase is a liquid. [64] Present day liquid chromatography is referred to as high performance liquid chromatography (HPLC) because the method utilizes

very small packing particles at a high pressure for increased separation performance. In HPLC the sample is forced by a liquid at a high pressure through a column that is packed with a sorbent that acts as a stationary phase leading to separation of the sample. The sorbent is typically a granular material made of solid particles such as silica or various polymers. The components of the sample are separated from each other due to their different degrees of interaction with the sorbent particles. The pressurized liquid is typically a mixture of solvents is commonly referred to as the mobile phase. Methods in which the stationary phase is more polar than the mobile phase are termed normal phase liquid chromatography and the opposite is termed reversed phase liquid chromatography. [64]

Just like in the case of gas chromatography, liquid chromatography can be coupled to a mass spectrometer to analyze the separated mixtures. However, different ionization methods must be utilized to ionize samples for analysis by mass spectrometers after HPLC. The most commonly used ionization methods are thermospray and electrospray. Thermospray may be defined as the controlled partial vaporization of a liquid as it flows through a heated capillary tube. Upon exiting the heated capillary, the rapidly expanding sample vapor converts the remaining liquid stream to an aerosol. Electrospray is a method of generating a very fine liquid aerosol through electrostatic charging. In electrospray, a liquid is passed through a nozzle which has a potential difference applied to it. This generates an aerospray of droplets by electrically charging the liquid to a very high voltage. The charged liquid in the nozzle becomes unstable as it is forced to hold more and more charge. The liquid reaches a critical point, at which no more charge can be contained in the droplets. The liquid is released as a cloud of tiny fine droplets. These droplets can then by analyzed by mass spectrometry. Common applications of LCMS are Pharmacokinetics, proteomics, the analysis of carbohydrates and drug development. [64]

3.4 Monosaccharide Composition Analysis

The first step in characterizing the CPS structure is determining which monosaccharide structures are present in the sample. Analysis is performed by GCMS. However, native sugars are not volatile due to hydrogen bonding and must be derivatized into volatile compounds. The Monteiro lab uses alditol acetates as volatile derivatives of carbohydrates that can be analyzed by GCMS. In a series of steps, the carbohydrates undergo hydrolysis, reduction and acetylation to produce the alditol acetate derivative. [65] The hydroxyl groups on the carbohydrates are acetylated to severely lowering the boiling point and producing a volatile derivative.

The derivatization begins with acid hydrolysis, as shown in the first step Figure 21, where a small amount of the polysaccharide (0.5 – 1.5 mg) is dissolved in 1 mL of 4M trifluoroacetic acid (TFA) at 105°C for 4.5 hours. The harsh acidic conditions break down the glycosidic linkages of the polysaccharide, breaking the polysaccharide down into its monosaccharide units. TFA is used for the procedure as it is easily evaporated compared to other acids due to high volatility.

Once the TFA has been removed by evaporation, 1.5 mL of water is added and, the sample is reduced with a scoop of sodium borodeuteride ($NaBD_4$) and the reaction is left to occur overnight. This is shown in the second step of Figure 21. Reduction ensures that all monosaccharides are reduced to their open chain alditol forms. If the monosaccharides would persist in their ring forms the analysis via GCMS would be very troublesome as each sugar would be able to produce up to 5 peaks on the chromatogram due to having the possibilities of α/β pyranose/furanose ring structures. Having monosaccharides reduced to alditol structures ensures there will be one peak per monosaccharide residue on the chromatogram for simple

analysis. Sodium borodeuteride is chosen over sodium borohydride as the reducing agent because C1 of the alditol will be reduced with a deuterium atom instead of a hydrogen atom. The top of the alditol is now one Dalton heavier which distinguishes the top and the bottom of the molecule. [65]

Reduction is stopped by adding glacial acetic acid to the mixture. Excess sodium borohydride decomposes to sodium borate in the presence of acetic acid. Borate compounds form covalent complexes with diols which inhibit acetylation reactions. The removal of these borate compounds is accomplished by adding 3mL of a 5% acetic acid (AcOH) and 95% methanol (MeOH) solution which forms tetramethyl borate gas which can be evaporated away. The evaporation leaves behind sodium acetate which serves as a catalyst for the acetylation reaction. Acetylation is performed by adding 2mL of acetic anhydride (AcOAc) and heating at 105°C for 90 minutes as shown in the final step of Figure 21. This step acetylates all the free hydroxyl groups on the alditol, removing all possibilities for hydrogen bonding and producing a volatile alditol acetate that can be analyzed via GCMS. The alditol acetate product is then extracted using dichloromethane (DCM) and then dried by a sodium sulphate column. The purified derivate is then evaporated to dryness and dissolved once more by DCM for analysis by GCMS.

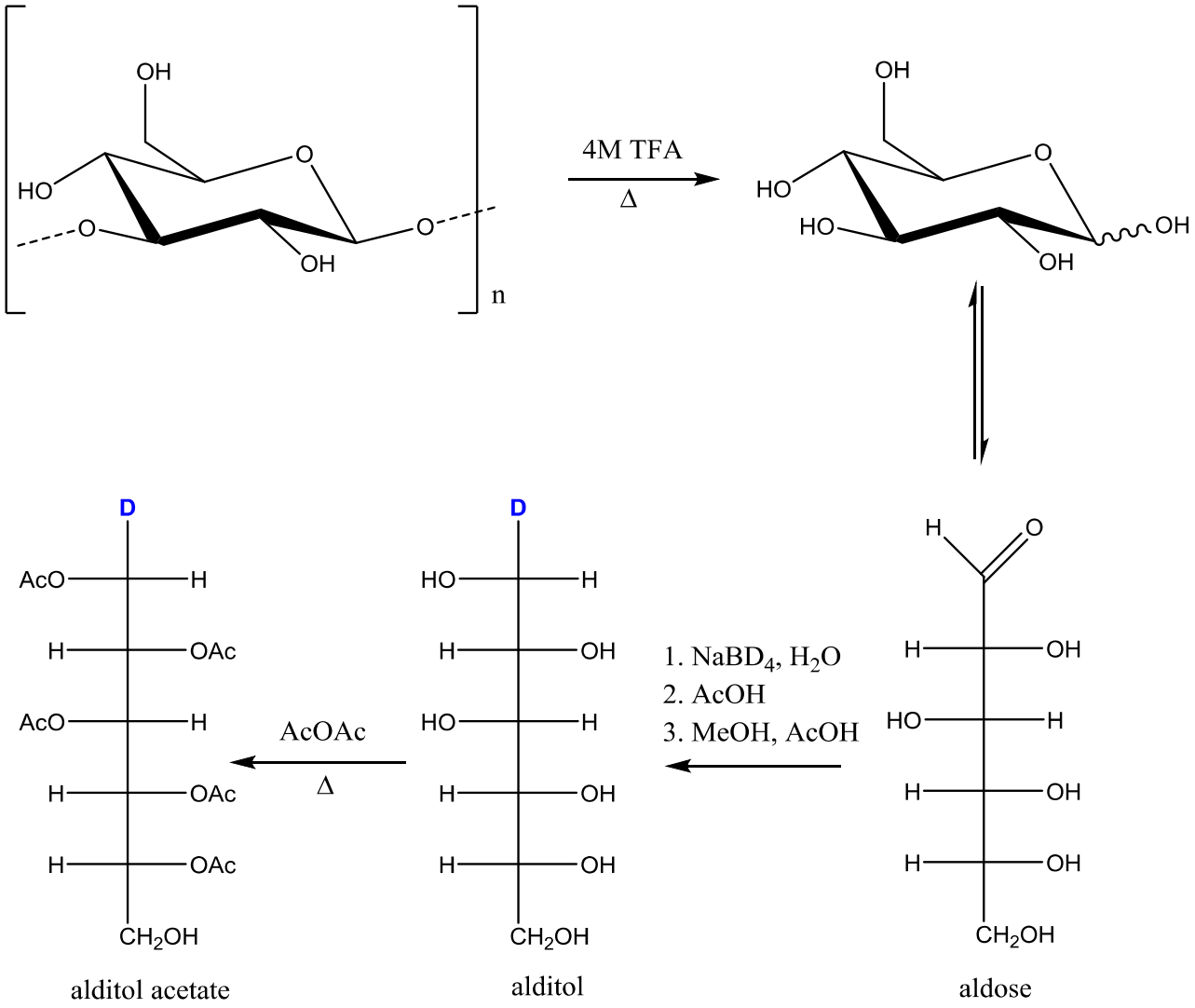


Figure 21: The alditol acetate method. The polysaccharide is hydrolyzed to its monosaccharides with 4M trifluoroacetic acid (TFA). Reduction with sodium borodeuteride (NaBD4) reduces the sugar to its alditol form and the acetic anhydride (AcOAc) acetylates the hydroxyl groups to produce an alditol acetate.

3.5 Monosaccharide Linkage Analysis

Alditol acetates allow us to determine which monosaccharides the capsular polysaccharide is comprised of. To see how these monosaccharides are linked a new derivative must be synthesized. The derivative is known as a partially methylated alditol acetate (PMAA).

Production of the derivative starts with methylation before conversion to alditol acetates. Methylation is done to methylate hydroxyl groups that are not involved in linkages. When the polysaccharide is broken down into monosaccharides via hydrolysis, the hydroxyl groups that are involved in linkages remain as hydroxyl groups. After going through the alditol acetate procedure they end up becoming an acetyl group. Methylated hydroxyl groups remain labelled with the methyl group at the end of the procedure allowing for a very simple distinction between hydroxyl groups involved in linkages.

Methylation is performed by dissolving the sample in 3mL of DMSO and treating the sample with a large excess (2.2 mL) of iodomethane (CH_3I) and a scoop of powdered sodium hydroxide. This is shown in the first step of Figure 22. The mixture is left to react for 5 hours to ensure complete methylation of the hydroxyl groups. The methylated polysaccharide is purified by transferring the mixture to a vial of 50:50 DCM/water mixture (~2mL). After mixing, the sample is centrifuged to ensure good separation of the layers. The water layer is discarded and the DCM layer is washed with water two more times to ensure \purification. In the end, the DCM is evaporated to attain the methylated polysaccharide. The polysaccharide now undergoes the same derivatization procedure as an alditol acetate to become a partially methylated alditol acetate.

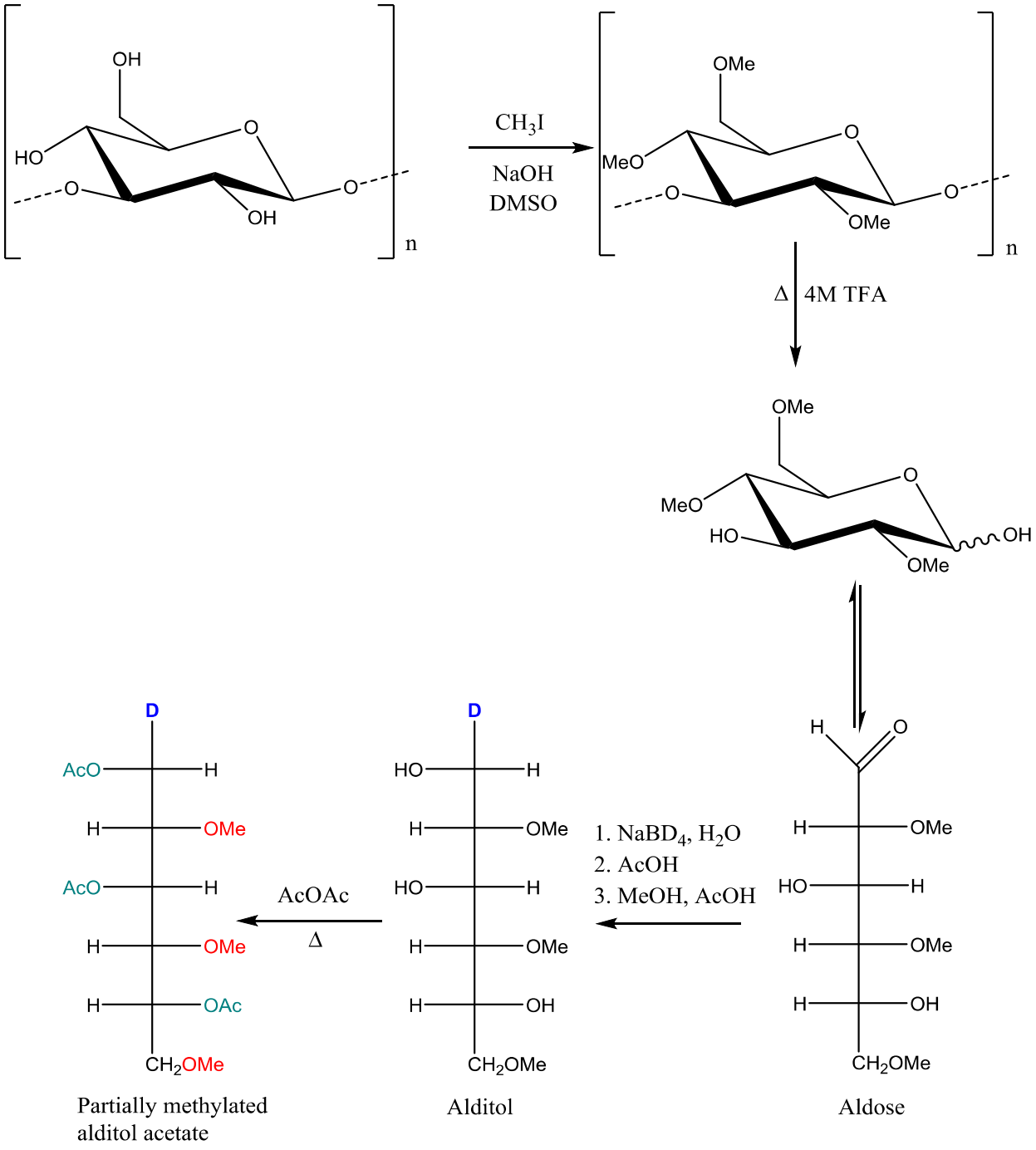


Figure 22: The partially methylated alditol acetate method. B-D-Glucose undergoes methylation using iodomethane. The polysaccharide is then hydrolyzed to its monosaccharide with 4M trifluoroacetic acid (TFA). Reduction with sodium borodeuteride (NaBD4) reduces the sugar to its alditol form and the acetic anhydride (AcOAc) acetylates the hydroxyl groups to produce a partially methylated alditol acetate

3.6 Fragmentation of Alditol Acetates

EI produces a radical cation that is susceptible to fragmentation to more stable ions. Primary fragmentation occurs first and results in the cleavage of the carbon carbon bonds in the main chain of the alditol acetate. Primary fragmentation is shown in Figure 23. For permethylated alditol acetates, the primary fragments prefer to cleave near methoxylated carbons with a large preference for vicinal methoxylated carbons. Carbons that have acetoxy groups attached can be cleaved but the occurance is much lower due to the unfavourable electron-withdrawing effect of the carbonyl group. [40]

After the ion has undergone primary fragmentation, secondary fragmentation occurs. This is displayed in figure 24. For alditol acetates common secondary fragments are: acetic acid (-60 m/z), ketene (-42 m/z) and formaldehyde (-30 m/z). For permethylated alditol acetates, as there are methylated and acetylated carboxy grops present, the loss of methanol (-32 m/z) is also very commonly seen when looking at the fragmentation patterns.

For linkage analysis the intensity of the fragment peaks is a very useful indicator for figuring out common linkages for the analyte. Stable primary fragments that can be detected by GCMS form around methoxylated groups and not the acetoxylated groups. The positive charge that is formed is more stable on the methoxylated carbon atom than the acetoxylated carbon. Reduction with sodium borodeuteride allows for identification between the top and the bottom of the molecule. The top of the molecule will have a higher mass per charge by 1 Dalton than the bottom of the molecule.

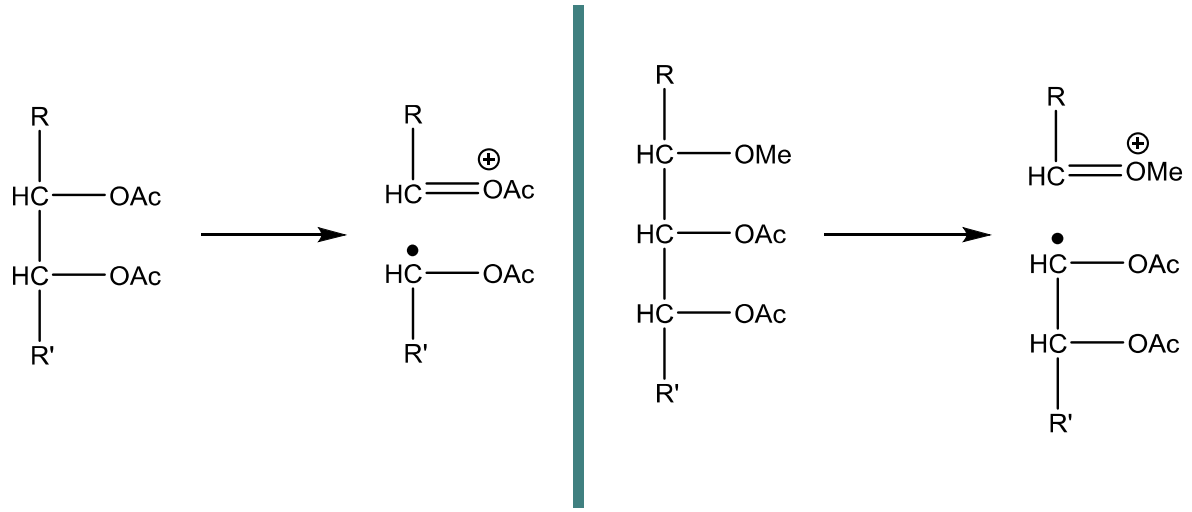


Figure 23: Primary fragmentation of alditol acetates and partially methylated alditol acetates. Fragmentation occurs at methoxy groups due to the unfavourable electron withdrawing effect of the carbonyl groups.

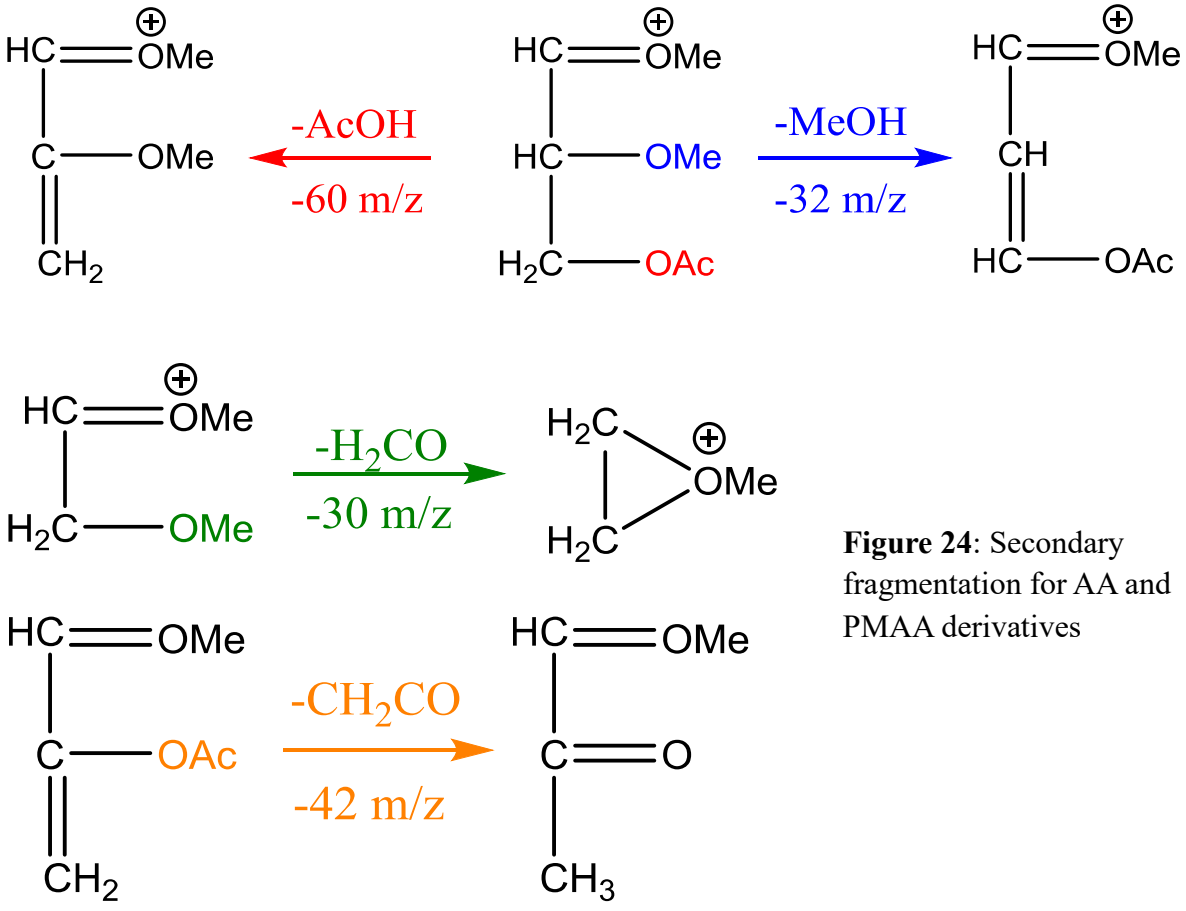


Figure 24: Secondary fragmentation for AA and PMAA derivatives

3.7 Instrumentation

The composition of the polysaccharide derivatives is analyzed using gas chromatography and gas chromatography-mass spectrometry. The derivatives have been synthesized as heat stable and volatile compounds to allow for analysis by GC and GCMS.

Gas chromatography is performed using a Varian 3400 instrument with a 30m x 0.25mm DB-17 fused silica column with a film thickness of 0.15 µm. 1 µm of sample is injected for analysis using helium as the carrier gas; the sample is detected by a flame ionization detector and the results are tracked by the program Peaksimple. Gas chromatography is performed to obtain an accurate ratio of the components of the analyte. Derivatizing the sugars into alditol acetates produces one peak per monosaccharide component and the intensity of the peaks provide an accurate representation of how much of the component is present. Identification is performed by comparing relative retention times of the sample to relative retention times collected for standards for many different monosaccharides.

To ensure that the results are correct GC-MS analysis is performed using 70 eV electron impact (EI) ionization on a Thermo Polaris Q with an Agilent DB-17 GC column. Each monosaccharide produces a unique mass spectrum, which can be used like a fingerprint, as each residue has a specific fragmentation pattern. In Figure 25, epimers, such as Glucose, Galactose and Mannose, produce the same fragmentation pattern however the spectra can be distinguished by different intensities of produced fragments.

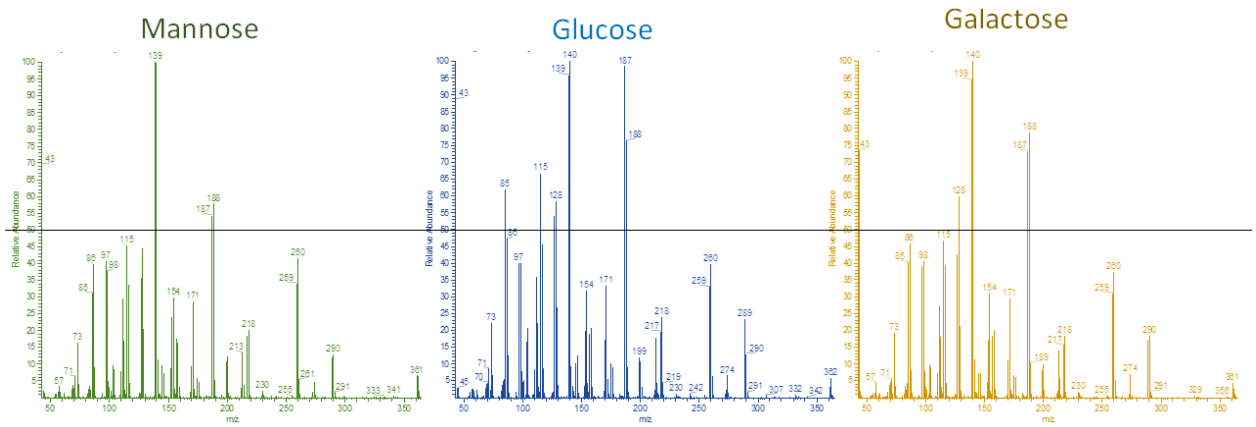


Figure 25: Mass spectra of Mannose, Glucose and Galactose.

LC-MS analyses were performed in the Mass Spectrometry Facility within the Advanced Analysis Centre at the University of Guelph. Samples were injected into a Dionex UHPLC UltiMate 3000 liquid chromatograph interfaced to an amaZon SL ion trap mass spectrometer (Bruker Daltonics, Billerica, MA). A C-18 column (Agilent Poroshell 120 2.7 micron particle size, 150 mm x 4.6 mm) was used for chromatographic separation. The initial mobile phase conditions were 98 % water (20 mM ammonium acetate) and 2% acetonitrile. The gradient went to 98% acetonitrile in 30 min. The flow rate was maintained at 0.4 mL/min. The mass spectrometer electrospray capillary voltage was maintained at 4.5 kV and the drying temperature at 220 °C with a flow rate of 10 L/min. Nebulizer pressure was 40 psi. Nitrogen was used as both nebulizing and drying gas, helium was used as collision gas at 60 psi. The mass-to-charge ratio was scanned across the m/z range 50–2000 in enhanced resolution negative-ion auto MS/MS mode. The Smart Parameter Setting (SPS) was used to automatically optimize the trap drive level for precursor ions. The instrument was externally calibrated with the ESI TuneMix (Agilent). UV monitoring was at 210 nm.

3.8 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is used to determine structural information of the CPS. NMR spectroscopy depends on the interaction of magnetic nuclei of with a static magnetic field. Magnetically active nuclei must have non-zero spins. Spin of the nucleus is determined by the composition of the subatomic particles that make up the nucleus. When there are an even number of protons and neutrons in the nucleus, the net spin of the nucleus is 0 and the nucleus is not magnetically active. When the nucleus is composed of an odd number of neutrons and protons the nucleus will have an integer spin and it will be magnetically active. If the sum of neutrons and protons is an odd number the nucleus will have a half integer spin and will also be magnetically active. The nuclei of isotopes: ^{1}H,^{13}C, ^{19}F and ^{31}P are all magnetically active with a spin of ½ and can therefore be analyzed using NMR spectroscopy.

The magnetically active nuclei generate magnetic moments that are proportional to the spin of the nuclei. Normally, the nuclear magnetic fields are randomly oriented. However, when the nuclei are placed in a magnetic field the nuclei align with or against the magnetic field. The energy difference between the two spin states is dependent on the external field strength.

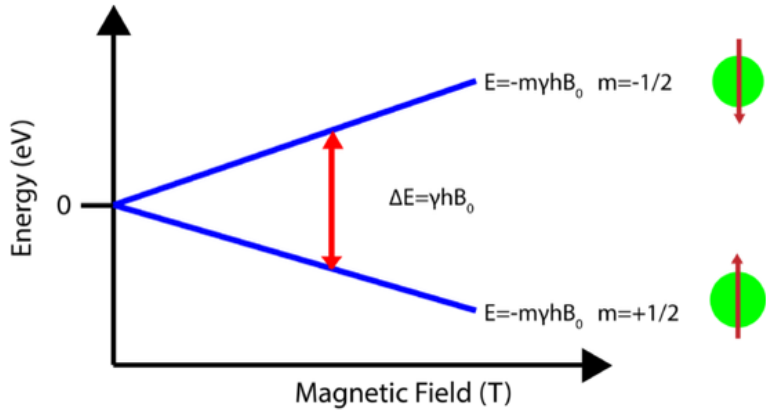


Figure 26: The effect of a magnetic field on magnetic nuclei. The nuclear energy levels undergo Zeeman splitting due to the presence of the magnetic field

For an NMR experiment to occur, an energy transition between the energy levels occurs. The energy transition will only occur when electromagnetic radiation of the correct frequency is applied to match the energy difference between the nuclear spin levels. Frequencies of these excitation magnetic fields typically fall in the radio frequency (RF) range of the electromagnetic spectrum. When a system is hit with the correct excitation frequency transitions will be induced not only from lower energy level to higher energy level, but there will also be downward transitions from the higher level to the lower level. Transitions result in absorption or release of energy. The probability of transitions occurring in either direction is equal, meaning that if the energy levels have the same population, the number of transitions will be the same and the interactions will cancel one another out. Only when the populations are not equal, a net absorption or release of energy that can be observed. In most cases, there are slightly more magnetic moments that align with the field than against it. Therefore, when excitation occurs there is a net release of energy. This can be visualized as a bulk magnetization vector that aligns with the direction of the magnetic field. When an RF pulse is applied along the x' axis the bulk magnetization vector rotates about the x' axis onto the y' axis. The extent of the rotation is determined by the duration of the pulse. When the pulse ends, the nuclei relax back to their equilibrium positions releasing a decaying. This decaying signal is picked up in a coil as an electromagnetic frequency that contains the sum of all the frequencies from the target nuclei. The signal is processed as a Free Induction Decay (FID)

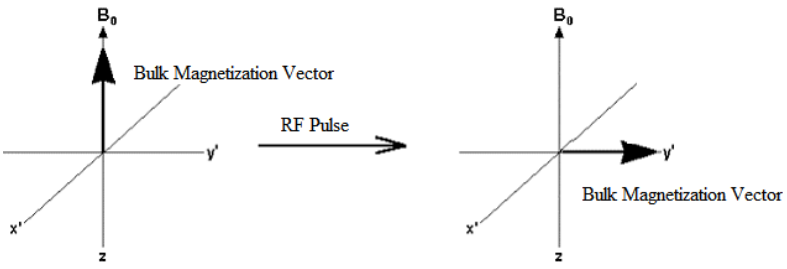


Figure 27: Application of the RF Pulse along the x' axis to the bulk magnetization vector. This effect is called transverse magnetization.

The power of NMR is based on nuclear shielding, which allows the method to be used for structural determination. Nuclear shielding occurs due to electrons orbiting nuclei. Electrons are also charged particles that generate their own magnetic field. The electron magnetic field runs anti-parallel to the external magnetic field. This magnetic field interferes with the applied external field and reduces the magnetic moment affecting the nuclei. This means that nuclei in different chemical environments will need different resonance frequencies to cause energy transitions. The shielding of nuclei allows for chemically different environments to be determined by performing a Fourier transform on the FID.

3.9 1D NMR Experiments

There are various 1D and 2D experiments that allow for the polysaccharide structure determination. The NMR experiments were performed on a Bruker AVANCE III 400 MHz spectrometer at 273K equipped with a cryoprobe. All data was analyzed using Topspin 2.1 and 3.1 software. All samples were prepared using 3 exchanges with deuterated water (D_2O) with freezing and lyophilisation occurring between each exchange. The deuterated samples were dissolved in 600 µL of D_2O for introduction to an NMR analysis tube. 3-Trimethylsilyl-tetradeutero sodium propionate (TSP) is used as the proton standard, with a δ_H 0ppm.

1D proton NMR (^{1}H-NMR) is the first experiment performed. The experiment targets all the hydrogen atoms present on the carbon atoms of a sugar. The proton NMR experiment has a very simple pulse sequence.

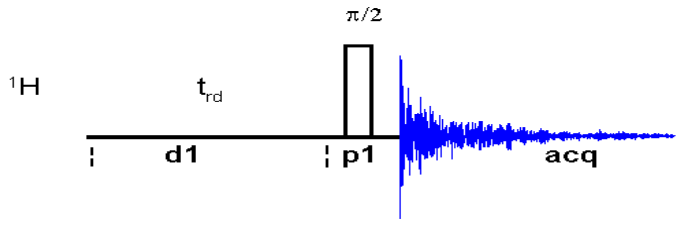


Figure 28: Pulse sequence for 1D proton experiment.

D1 is the interpulse delay and allows for the spin systems to relax back to equilibrium and align themselves with the applied magnetic field. For protons this relaxation period is about 2-4 seconds. The 90° RF pulse is applied at P1 and lasts in the 5-15 second microsecond range. This pulse excites the spin systems and when the systems decay back to equilibrium the signal is collected as a FID. The FID undergoes a Fourier transform to become an NMR spectrum. [66]

Proton NMR is useful when dealing with polysaccharides as the experiment will indicate the number of anomeric protons present in the polysaccharide. The anomeric protons will appear on the spectrum in the 4.4-5.5 ppm region. The alpha-anomer resonates downfield from the beta anomer. [67] One has to be careful in this region as the HOD peak from the D_2O solvent can interfere with anomeric signals. The number of anomeric protons observed will give insight to the number of monosaccharide residues in the structure of the monosaccharide. Another characteristic region is found from 1.6 to 1.9 ppm where signals from deoxy- functional groups are observed. [67] The ring protons that are not in the anomeric position are most commonly found in the region of 3.0 to 4.3 ppm. The ring region suffers from heavy overlap due to all the similar chemical environments for these protons. There is a lot of ambiguity in assigning these protons and other 1D and 2D experiments must be applied.

1D phosphorus NMR (^{31}P-NMR) is much less sensitive than Proton NMR but the experiment has its uses when dealing with surface carbohydrates. In the case of many surface polysaccharides, the common phosphorus containing compounds are: O-Methyl phosphoramidate and phosphate bridges that are used to link repeating units together. The O-methyl phosphoramidate region on a ^{31}P-NMR spectrum is found at 14-15 ppm. The phosphate region is found at 0 ppm. These regions are useful for analysis when purifying the polysaccharide. The core polysaccharides contain many phosphate groups that must be purified

away when working with these various surface polysaccharides. It is possible to track the purity of a polysaccharide looking at the intensity of the phosphate peaks. However, some CPS structures contain phosphate groups crucial to the function of the surface polysaccharide.

3.10 2D NMR Experiments

In one-dimensional NMR the signal is recorded as one time variable and then Fourier transformed to give a spectrum which is a function of one frequency variable. In two-dimensional NMR the signal is a function of two time variables, t1 and t2, and the data must therefore be Fourier transformed twice to yield a spectrum which is a function of two frequency variables. The pulse sequence for 2D-NMR experiments is shown below in figure 28. [66]

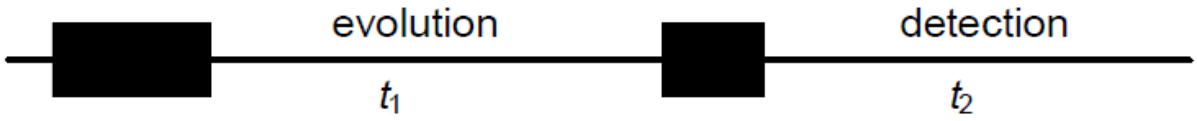


Figure 29: General pulse sequence for 2D NMR experiments

The sample is first excited by one or more pulses. This pulse generates states that are set up to interact in a useful way. The resulting magnetization is allowed to evolve for t1 which is known as the evolution period. The evolution period labels the magnetization as a function of t1. After the evolution period, a second period of pulses is applied known as the mixing time. The mixing time is responsible for transferring the magnetization through bonds. The signal is now ready to be detected during t2. Magnetization that did not get transferred will appear in the same frequency at t1 along the diagonal. Magnetization that was transferred during the mixing period will appear at a frequency off the diagonal. These off diagonal signals are referred to as cross peaks. These cross peaks indicate the coupling between two nuclei. During the experiment, t1 is incremented, the evolution time starts close to zero for the first spectrum and is then increased

from one run to the other. Fourier transformation of the t2 time domain gives us a set of spectra as a function of t1 step size. A second Fourier transformation of t1 gives a 2D spectrum revealing the frequencies of modulation that were occurring during t1.

Correlation spectroscopy (COSY) is a 2D-NMR experiment that identifies spin-spin coupling within molecules. The experiment is most useful for determining proton-proton correlation that exists within molecules. The pulse sequence is shown below in figure 29. [66]

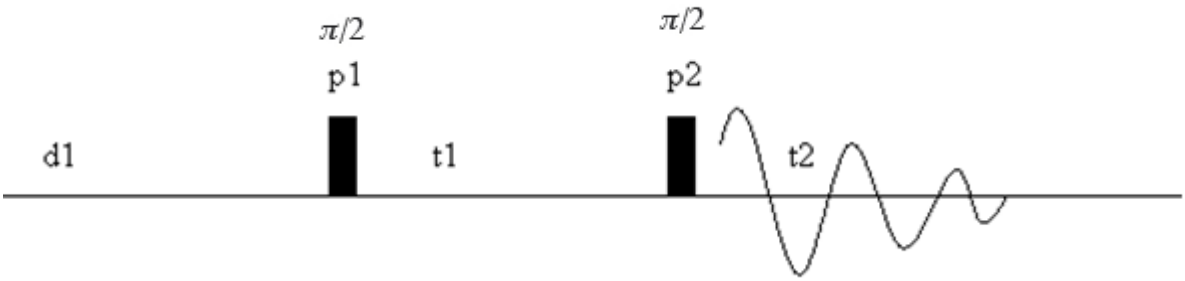


Figure 30: Pulse sequence for correlation spectroscopy (COSY)

The experiment once begins with an interpulse delay that allows for the spin systems to relax back to equilibrium. A 90° pulse is applied that generates transverse magnetization which rotates the bulk magnetization vector from the z axis to the y axis. [66]

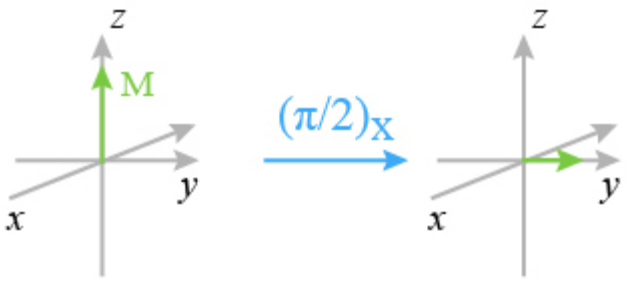


Figure 31: Transverse magnetization

During the relaxation period t1 the spin system precess between the x and y axis due to random fluctuations in the magnetic field.

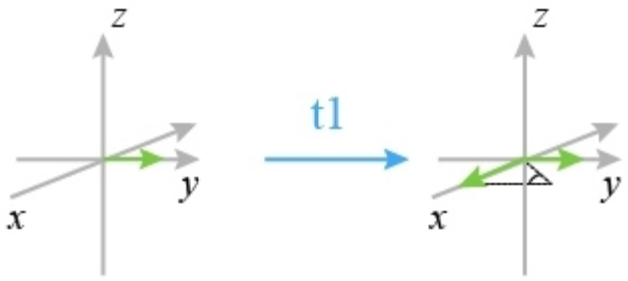


Figure 32: Spin precession

The second 90° pulse is now applied during the mixing time. This rotates the y component in the negative z direction

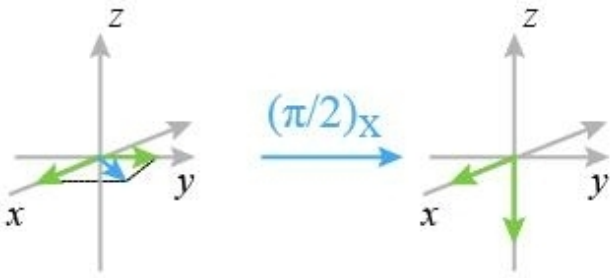


Figure 33: Effect of Pulse 2

The second pulse induces magnetization transfer. an exchange of magnetization occurs between all the nuclei that are coupled together. If the magnetization is not transferred the frequency remains on the diagonal. Coupled nuclei are spotted by the presence of cross peaks outside the diagonal. The coupling of protons is limited with to a 3 bond length. [66] A single proton can generate a singlet, doublet, triplet, quartet, doublet of doublets or a multiplet depending on its chemical environment. [68] Starting from an easily identifiable proton, such as

the anomeric proton, all the protons on the polysaccharide can be assigned. Some ambiguities may still arise and other NMR experiments can be carried out to solve these ambiguities.

Heteronuclear single-quantum correlation spectroscopy (HSQC) is a 2D-NMR experiment that detects correlations between two different nuclei which are separated by one bond. This method results in one cross peak per pair of coupled nuclei. The pulse sequence for an HSQC experiment is shown below.

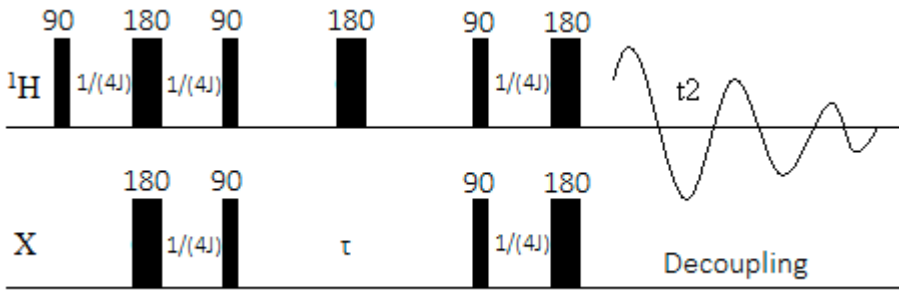


Figure 34: HSQC Pulse Sequence

The HSQC experiment is dependent on the Insensitive Nuclei Enhanced by Polarisation Transfer (INEPT) pulse sequence. [66] The INEPT pulse sequence is used to transfer magnetisation from a sensitive nucleus (normally ^{1}H) to a less sensitive nucleus via scalar coupling. The INEPT pulse applies a 90° pulse the proton spin systems to create transverse magnetization. There is no an evolution delay that allows the magnetization to evolve. Simultaneous 180° ^{1}H and X pulses are applied during the middle of this evolution period in order to remove ^{1}H chemical shift evolution. A simultaneous 90° ^{1}H and X pulse is now applied to transfer the polarization from ^{1}H to X. This magnetization is allowed to evolve during the τ evolution period. Heteronuclear 1H-X couplings are refocused by applying a 180° ^{1}H pulse at the middle of this period. The INEPT pulse sequence must be reversed to reconvert the antiphase magnetization of the low sensitivity X nuclei to inphase magnetization for the ^{1}H nuclei. This is done by applying a simultaneous 90° ^{1}H and X pulse to transfer polarization from X to ^{1}H to

achieve antiphase ^{1}H magnetization with respect to X. There is now another evolution period that allows for the evolution of the heteronuclear coupling constants. Simultaneous 180º ^{1}H and X pulses are applied in the middle of this period to refocus the ^{1}H magnetization. This magnetization gives the desired signal signal after the t2 evolution period. The decoupled signals are removed by running the experiment twice with the phase of one pulse reversed so that the signs of the undesired peaks are reversed and the undesired decoupled peaks will cancel out. [66]

The most common atoms studied as X are ^{13}C and ^{15}N. For polysaccharides the ^{1}H-^{13}C experiments are the most relevant. The experiment allows for accurate determination of H1 and H6/H6' by the characteristic chemical shifts for the carbons these protons are attached to. The anomeric carbon is downfield at around 100 ppm and the H6 carbon is found at around 20 ppm. HSQC is the easiest way of finding out how many monosaccharides are in a polysaccharide. [68]

Heteronuclear multiple bond correlation (HMBC) is another 2D experiment that can detect coupling between 2 heteroatoms. However, the technique is useful as it can detect correlation over multiple bonds while HSQC can only detect correlation across 1 bond. The pulse sequence for the HMBC experiment is shown below in figure 34.

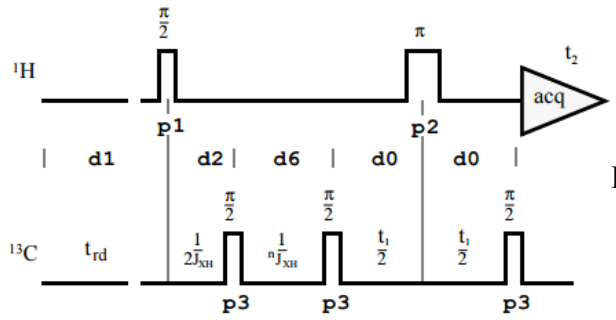


Figure 35: HMBC Pulse Sequence

The HMBC pulse sequence begins with a 90° pulse that is applied to the proton nuclei that creates transverse magnetization. This pulse is followed by a 90° pulse applied to the carbon

nuclei that is responsible for supressing all the one bond correlations in the 2D spectrum. The pulses are responsible for creating a heteronuclear multiple quantum coherence for ^{1}H's directly coupled to a ^{13}C nucleus. This coherence is unwanted as we are looking for multiple bond correlations and the unwanted coherence is removed from the spectrum by phase cycling the 90° pulse applied to the carbon nuclei. Removal of these signals is not necessary but it leads to a simpler final spectrum. [69] After a small delay, another 90° pulse is applied to the carbon nuclei which creates the desired heteronuclear multiple quantum coherence for 1H's coupled to a 13C nucleus 2 or 3 bonds away. This is followed by an evolution time t_1. Halfway through t_1 a 180° pulse is applied to the proton nuclei which refocuses the proton spins and only allows the chemical shifts to evolve. A final 90° degree pulse is applied to the ^{13}C nuclei which converts the double-quantum coherences back to antiphase magnetisation which is immediately detected. Phase cycling of this pulse removes all signals from proton nuclei that are not long range coupled to any carbon nuclei. [69] HMBC is useful when dealing with polysaccharides as the method can detect correlations across glycosidic linkages allowing for assignment of linkages between monosaccharides.

3.11 Instrumentation

Bruker 400 and 600 MHz NMR instruments were used to carry out the NMR experiments. Samples were dissolved in 99% D_2O and then freeze dried. This was repeated 3 times to ensure D_2O exchange. The final sample was suspended in 600 µL 99.9% D_2O. The Heteronuclear Single Quantum Spectroscopy (HSQC) involving ^{1}H and ^{13}C nuclei was performed using standard Bruker software pulse sequences, and performed at 22 °C with 3-(Trimethylsilyl)-Propionic acid (TSP) in D_2O as reference.

Chapter 4: Results and Discussions

4.1 Carbohydrate Composition of HS: 44 CPS

The GC profile of the extracted polysaccharide, which had been derivatized to alditol acetates, is shown in figure 37. The revealed monosaccharide components were glycerol, glucose, galactose, ribose as well as heptose constituents which were assigned as 6-deoxy-3-O-methyl-*altro*- heptose (6d-3-OMe-*altro*-Hep), 6-deoxy-*altro*-heptose (6d-altro-hep) and 6-deoxy-*galacto*-heptose (6d-*galacto*-Hep). Furanose configurations of the 6d-heptoses were confirmed by comparing the retention times to standards in lab. The existence of furanose configurations was later confirmed due to characteristic fragmentation patterns found from linkage analysis. Appearance of these 6d-oxy heptoses in *C. jejuni* serotype HS: 44 strain: 2871 confirms that the strain possesses the capability of synthesizing unusual 6-deoxy-heptoses containing unique ring configurations. In addition, the presence of these heptoses along with the presence of galactose, glucose and glycerol suggests that *C. jejuni* serotype HS: 44 generates two independent capsular polysaccharides in which one of the CPS contains the same monosaccharide composition as that of *C. jejuni* strains 856 (HS: 1) and 3087 (HS: 1,44). These capsules are composed of a galactose residue that is linked to a glycerol phosphate and two fructose residues modified with MeOPN.

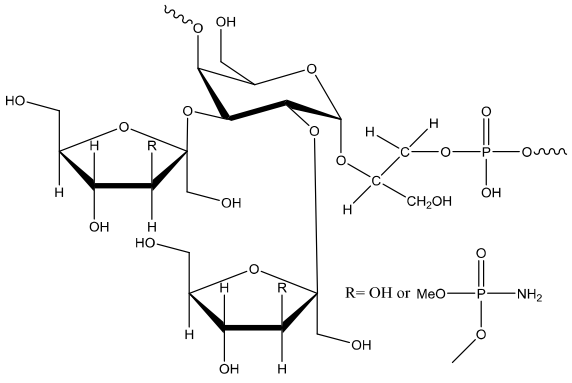


Figure 36: HS: 1 CPS structure

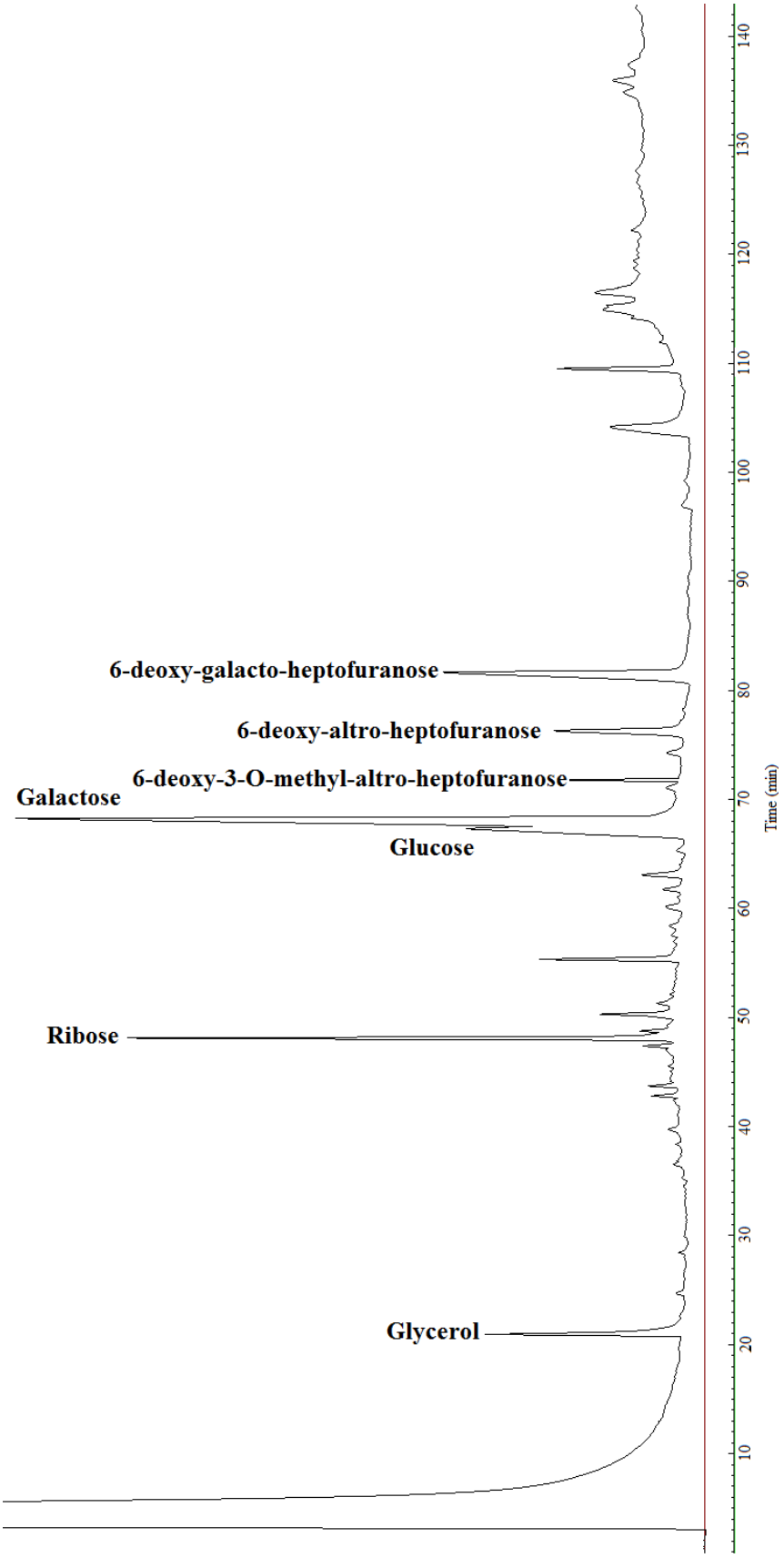


Figure 37: GC Profile of crude alditol acetate

4.2 NaOH Treatment of *C. jejuni* Capsular Polysaccharide

Once the crude capsular polysaccharide had been obtained from the core polysaccharide and identified, the goal of the research was to isolate the heptose-containing HS: 44 capsule from the HS: 1 capsule. This separation was attempted by performing a 10% NaOH treatment on the capsule for 7 days. The goal of the treatment was to cleave the phosphate linking bridges that are possibly responsible for linking the two capsules. The treatment was followed up by dialysis of the sample in a Spectrapor 1kDa molecular weight cut off (MWCO) dialysis membrane in a beaker of running deionized water for 7 days. The goal of the dialysis was to remove the HS: 1 capsule after it had been cleaved as a result of the NaOH treatment as the HS: 1 polysaccharide is smaller than the larger heptose-containing HS: 44 polysaccharide. The resulting polysaccharide was analyzed using GC-MS and NMR.

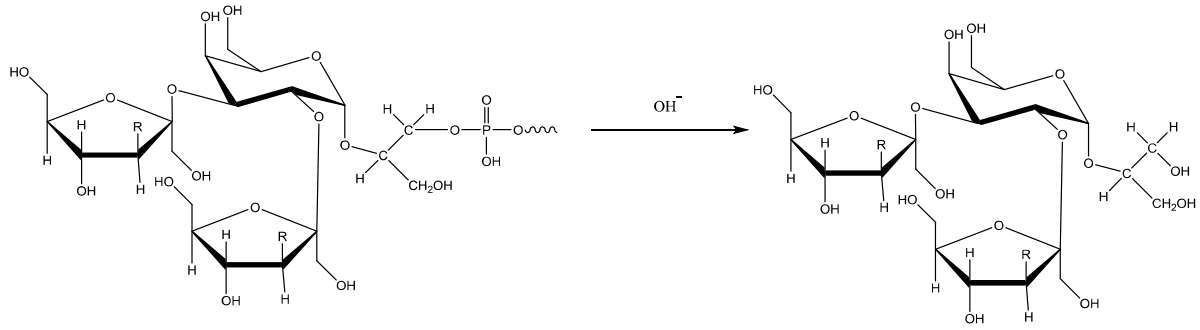


Figure 38: Proposed NaOH treatment that would cleave the phosphate bridge between the two capsules.

4.3 Alditol Acetates of NaOH Treated Capsular Polysaccharide

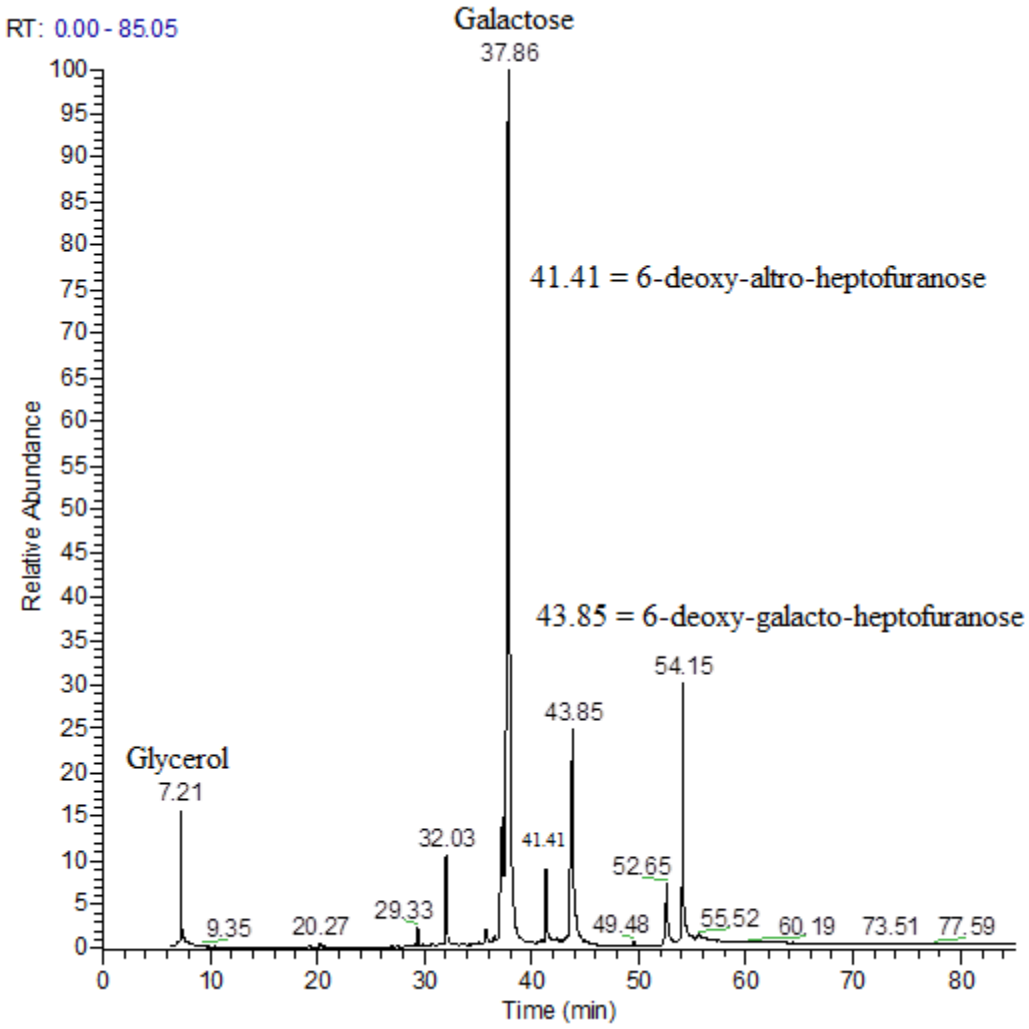


Figure 39: Gas chromatography profile from GCMS analysis of NaOH treated CPS of *C. jejuni*

Figure 39 above shows the effects of the NaOH treatment on the HS: 44 capsule. The ribose and 6-deoxy-3-O-methyl-altro-heptafuranose that were present in the crude sample have been removed. However, the galactose and glycerol from the HS: 1 capsule remained in the sample showing that the 2 capsules were not successfully separated by the NaOH treatment.

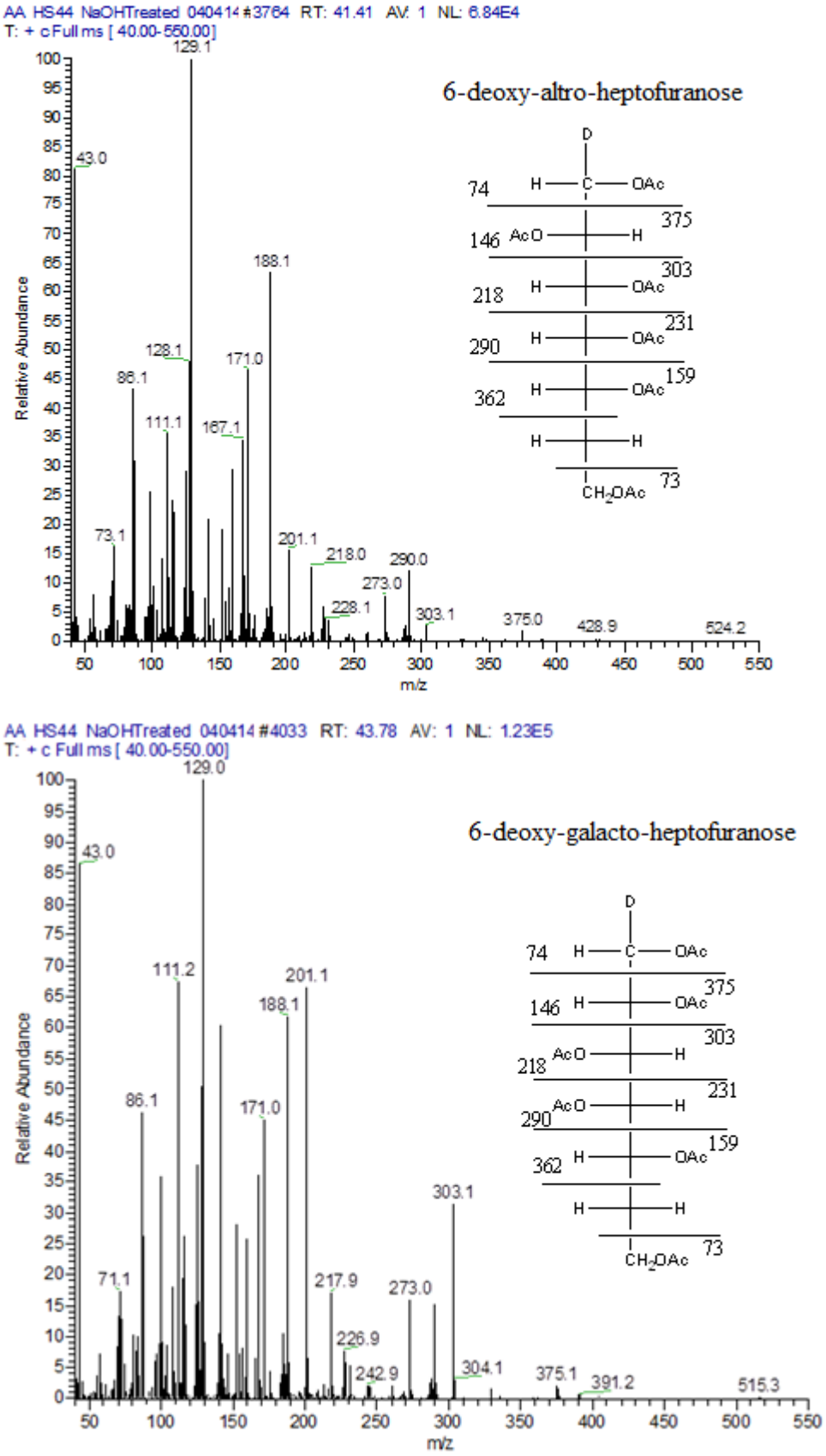


Figure 40. Mass spectra of the heptose constituents of the HS:44 capsule.

T

Figure 40 shows the EI-MS of 6-deoxy-heptitol acetates with *altro* and *galacto* ring configurations. The spectra of the two 6-deoxy-heptitol acetates generate the same fragments. The two heptoses can be identified by their preference for generating different amounts of characteristic fragments. Primary fragments observed at m/z 375, m/z 303, m/z 231, and m/z 159 confirmed the presence of a 6-deoxy-heptose. Other primary fragments were detected at m/z 73, m/z 74, m/z 218 and m/z 290. Fragmentation of primary fragments into secondary fragments can occur by the loss of acetic acid (m/z 60) and the loss of ketene (m/z 42). Loss of acetic acid produced fragments at m/z 99 and m/z 171 from primary fragments m/z 159 and m/z 231. Ketene fragmentation produced fragments m/z 129, from fragment m/z 171. A loss of an acetic acid followed by a consecutive loss of a ketene fragment produced the fragments: m/z 188, m/z, 201 m/z and 273 m/z, from m/z 290, m/z 303 and m/z 375. The fragment at m/z 141 was detected due to the loss of two acetic acid fragments and one ketene fragment from fragment m/z 303. These characteristic fragments were detected in other MS spectra of *C. jejuni* 6-deoxy heptoses which are found in pyranose conformations, such as strain CG8421 (HS: 23, 36), CG8486 (HS: 4, 13, 64) and BH-01-0142 (HS: 3, 13, 50). The fragments for the *galacto* ring conformation had a consistently higher intensity than fragments revealed from the *altro* ring. This was observed for the primary fragment at m/z, m/z 303 and secondary fragments at m/z 273, m/z 201 and m/z 141.

4.4 NMR Analysis

4.4.1 1D ^{1}H NMR

The ^{1}H NMR spectrum displayed in Figure 41 of the CPS of *C. jejuni* after NaOH treatment showed overlapping proton resonances in the anomeric region of the spectrum between δ 5.00 and 5.50 ppm. The resonances corresponded to galactose anomeric protons from HS: 1 and the heptose anomeric protons from HS:44. The signal at δ5.21 is the galactose from HS:1 lacking fructose substituents at the C-2 and C-3 positions. However, there are many anomeric peaks here that could correspond to many different α-glycosides which requires further analysis. There is also an overlapping ring proton region between δ 3.50 and δ 4.80 In addition; there is also a broad overlapping region that revealed methylene signals between δ 0.8 and δ 2.1. The region corresponded to the 6-deoxy moieties from the two 6-deoxy-heptose units which presence was consistent with the prior GC-MS results.

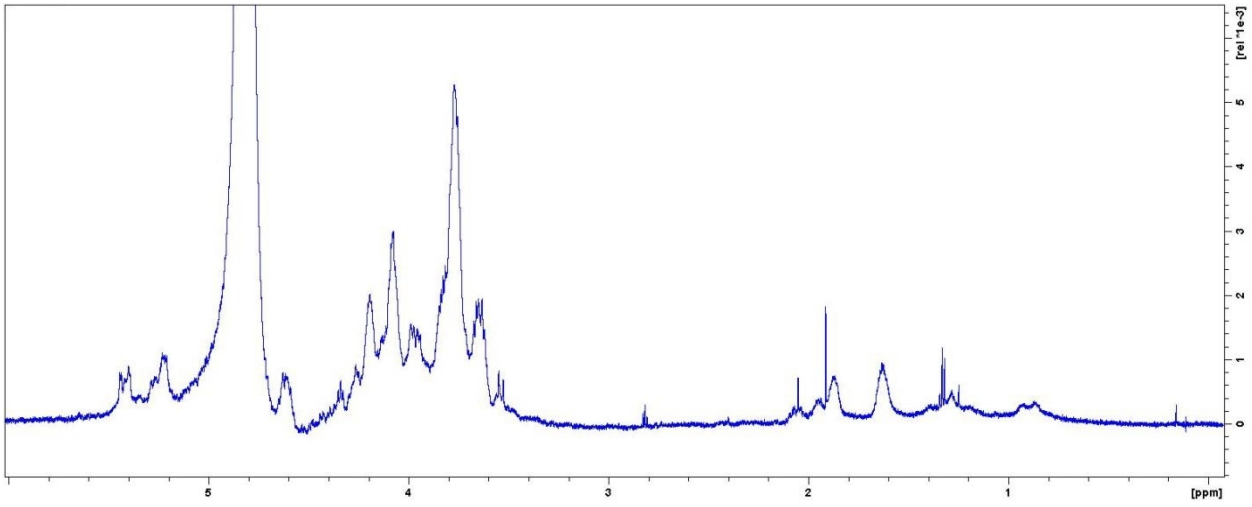


Figure 41: 1D ^{1}H NMR spectra of *C. jejuni* HS:44 CPS

4.4.2 1D ^{31}P NMR

1D ^{31}P NMR experiments were crucial in tracking the presence of O-methyl phosphoramidate and glycerol phosphate in the CPS over the course of purification and the

NaOH treatment. The spectra for before and after the treatment are shown in Figure 42. Before NaOH treatment, the CPS was found to contain both O-methyl phosphoramidate and glycerol phosphate species. The O-methyl phosphoramidate signal was located at δ 14.05 and various phosphate signals were located around δ 0. After NaOH treatment, the O-methyl phosphoramidate was cleaved from the CPS and was not detected anymore.

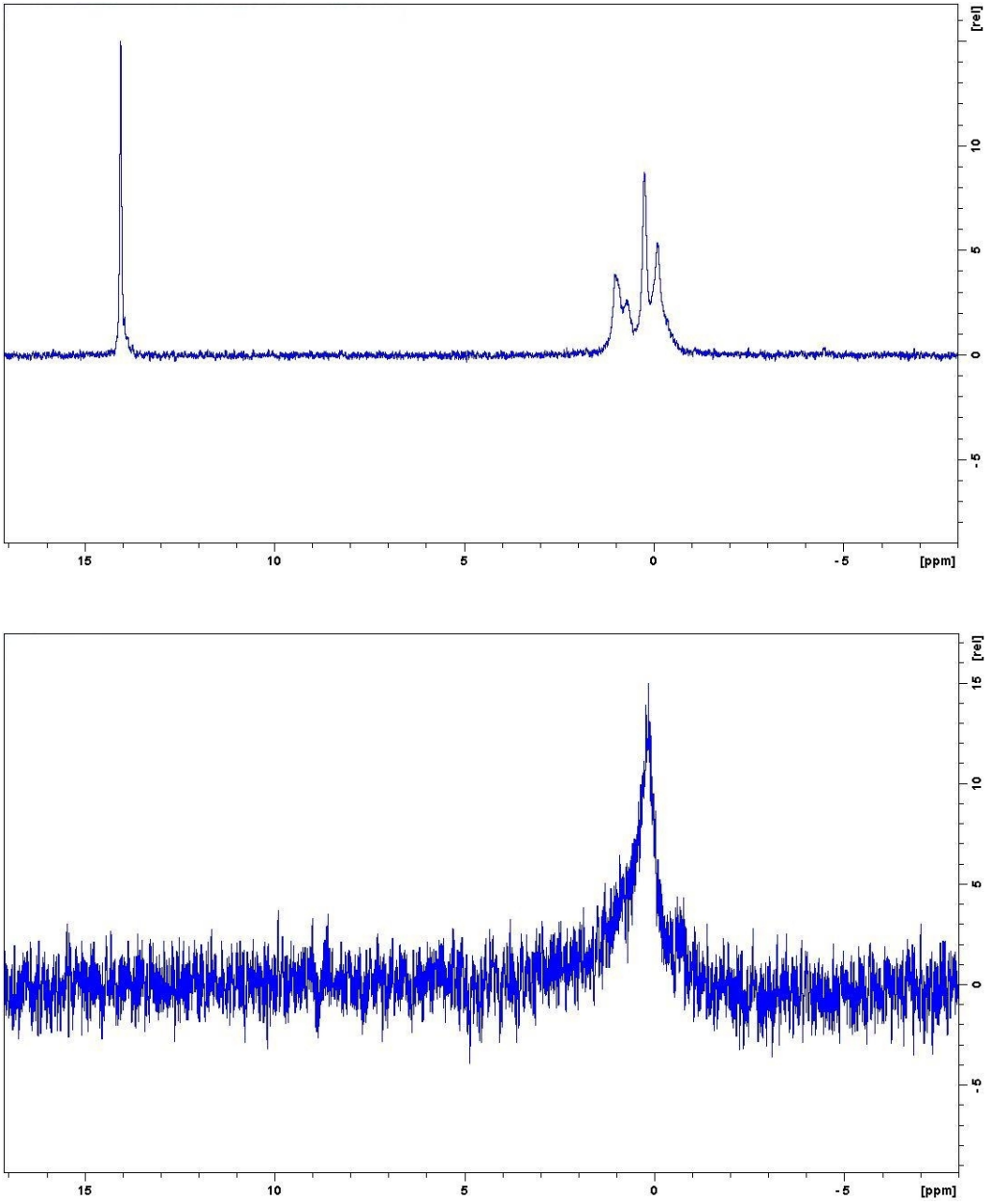


Figure 42: 1D ^{13}C HSQC NMR spectra of *C. jejuni* HS:44 CPS. The top spectrum is before NaOH treatment. The bottom spectrum is after NaOH treatment.

4.5 Discussion

The HS: 44 capsule, containing the HS: 1 capsule, was successfully purified from the cell wall of the *C. jejuni* cell mass. This was subject to NaOH treatment followed up by dialysis attempting to dialyze out the unwanted HS: 1 capsule so that the pure HS: 44 capsule could be obtained. After the dialysis ~25 grams of sample were obtained which is quite a large amount.

Collectively, the structural data obtained by GCMS and NMR analysis points to the conclusion that the week-long 10 percent base treatment was not successful in cleaving the two CPSs from one another. Instead, it seems that the NaOH treatment was simply responsible for cleaving the fructose units, which contained the phosphoramidate, from the HS: 1 CPS. If any phosphoramidate units were located on the heptoses of the HS: 44 CPS they would also be cleaved by this treatment. The NaOH treatment did not destroy the glycerol phosphate linkages as planned and another different treatment should be attempted in order to successfully separate the two CPS. A different separation method than dialysis might also be more successful in attempting to obtain the desired HS: 44 CPS. These methods could include running the sample through size exclusion chromatography column using P2 or G50 gel or even a more complicated separation technique such as ion exchange chromatography.

4.6 Presence of Mannose Units for *C. difficile*

When *C. difficile* was exposed to extended hot water-phenol treatment mannan units were detected in *C. difficile* biomass. Interestingly, five-day-old spore-rich *C. difficile* biomass preparations were observed to contain a greater concentration of mannan units. During the hot water-phenol extraction, the mannose rich material co-solubilized with other water soluble polysaccharides from *C. difficile* such as PS-I and PS-II leading to difficulties isolating the mannose rich material by size-exclusion chromatography. Nevertheless, enough material was obtained that allowed for analysis by gas chromatography-mass spectrometry and liquid chromatography-MS.

4.7 Sugar Compositional Analysis of *C. difficile*

Monosaccharide analysis detected greater than expected amounts of mannose when the crude aqueous preparation, which had been obtained from extended hot water-phenol extraction, was analyzed. Mannose was not expected to be found in the surface carbohydrate of *C.difficile*. Size exclusion chromatography was able to separate the mannose rich material from the water soluble *C. difficile* surface polysaccharides of PSI and PSII. In addition to greater than expected amounts of mannose, monosaccharide analysis of the isolated preparation revealed s glucose while no key structural markers of PSI and PSII (Rha of PS-I and GalNAc of PS-II) were observed in the chromatograms.

4.8 Linkage Analysis of *C. difficile*

GC-MS analysis of the permethylated alditol acetates derivatives revealed that the mannan linkages were composed of: end-groups [Man-(1→]; 2-monosubtituted linear units [→2)-Man-(1→]; and 2,6-disubstitued branch residues [→2,6)-Man-(1→]. Minor sugar linkage

types of 6-monosubstituted linear mannose [→6)-Man-(1→] and 3-monosubstituted linear mannose [→3)-Man-(1→] were also observed. A small amount of 4-linked glucose units were observed similar to those previously reported in Clostridia glucans.

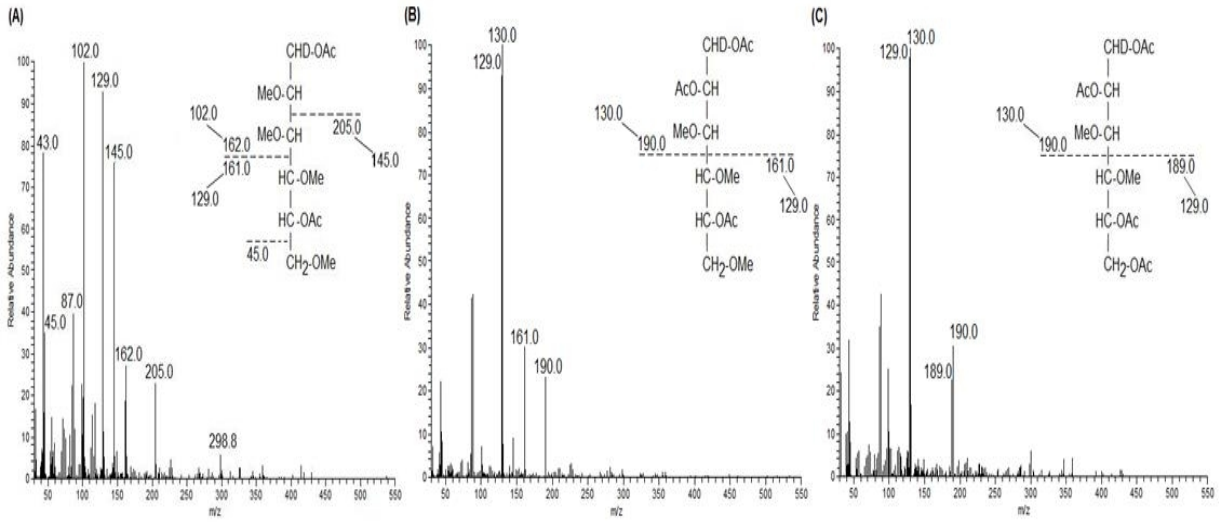


Figure 43. Mass spectra of the major Man linkage-types present in *C. difficile* mannan: (A) terminal Man; (B) 2-substituted Man; (C) 2,6-disubstituted Man.

Terminal mannan can be identified by the characterstic cleavage of the alditol acetate produced by the terminal mannan. The spectrum obtained is shown in figure 45 (A). The characteristic primary fragments that occur are 205 m/z, 162 m/z, 161 m/z and 45 m/z. These primary fragments are all formed from the cleavage of the carbon carbon bonds of the alditol acetate. From these primary fragments secondary fragments are formed by the loss of acetic acid (-60 m/z) and the loss of methanol (-32 m/z). For the terminal mannan the primary fragment of 205 m/z loses an acetic acid to produce a secondary fragment that weighs 145 Daltons. The 162 m/z primary fragment also loses an acetic acid to form the fragment found at 102 m/z. The primary fragment found at 161 m/z loses a methanol to produce the fragment found at 129 m/z. From these characteristic fragmentations the terminal mannose group can be identified when compared to a standard.

For the 2-substituted mannose there is a new characteristic fragmentation pattern. The PMAA has two primary fragments from cleavages at C3 from the top and bottom of the molecule. These cleavages produce the primary fragments found at 191 m/z and 161 m/z. The primary fragment at 190 m/z loses an acetic acid group to form the secondary fragment found at 130 m/z on the mass spectrum and the primary fragment at 161 m/z loses a methanol group to form the secondary fragment found at 129 m/z. Identification of these primary and secondary fragments allows for the identification of the 2-substitued mannose group.

The 2,6-disubstitued Man is identified by the presence of the 190 m/z , 189 m/z, 130 m/z and 129 m/z peaks on the mass spectrum. Primary fragments are represented by the peaks at 190 and 189 m/z and occur because of cleavages at C3 that occur from the bottom and top of the molecule. Both secondary fragments formed at 130 m/z and 129 m/z occur due to the loss of acetic acid by the primary fragments of the molecule.

4.9 Liquid Chromatography Mass Spectrometry

The presence of mannan moieties was also exposed by LC-MS, which yielded conclusive evidence for hexose-containing species of varying molecular weights. LC-MS fraction A consisted of several compositions in which the major species possessed molecular weights of 3395.41 Daltons and 3719.51 Daltons. Neighboring molecular ions are observed that differ by singular mannose extensions. The losses of 162 Dalton units signify the loss of one hexose residue. In our case, these are the losses of the mannose residues. It is apparent in both fractions that many losses of 162 occur. The LCMS spectra confirm that *C. difficile* contains these large chains of mannose in the structural polysaccharide network of the cell wall.

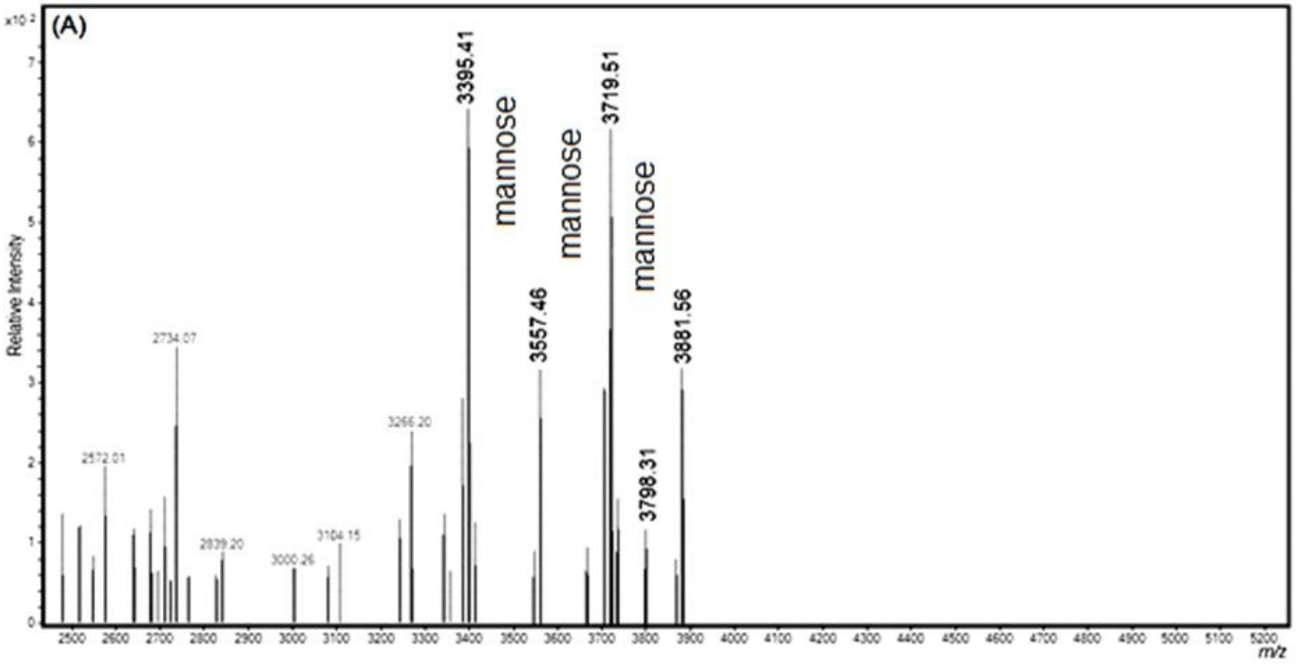


Figure 44: LCMS Fraction A

Here for fraction A it is observed that there is a fragment of 3881.56 Daltons. This fragment loses mannose molecules of molecular weight 162 three times to yield the fragment found at 3395.41 Daltons. These losses of 162 Daltons that occur many times confirm the presence of a long mannan chain in *C. difficile*.

A second fraction of the mannan containing material was analyzed by LC-MS. A similar fragmentation pattern with mannose-variable intensity patterns was observed for this fraction B (Figure 46), but in this case a species carrying an additional hexose was present at *m/z* 4044.63. This larger fragment meant that this fraction had one extra mannose residue on the mannan chain. This extra mannose residue meant there was one extra fragmentation of mannose from 4044.63 m/z to 3881.57 m/z. After this first fragmentation the remaining fragmentation pattern for the molecule is the same as the fragmentation pattern for the mannose residues in fraction A.

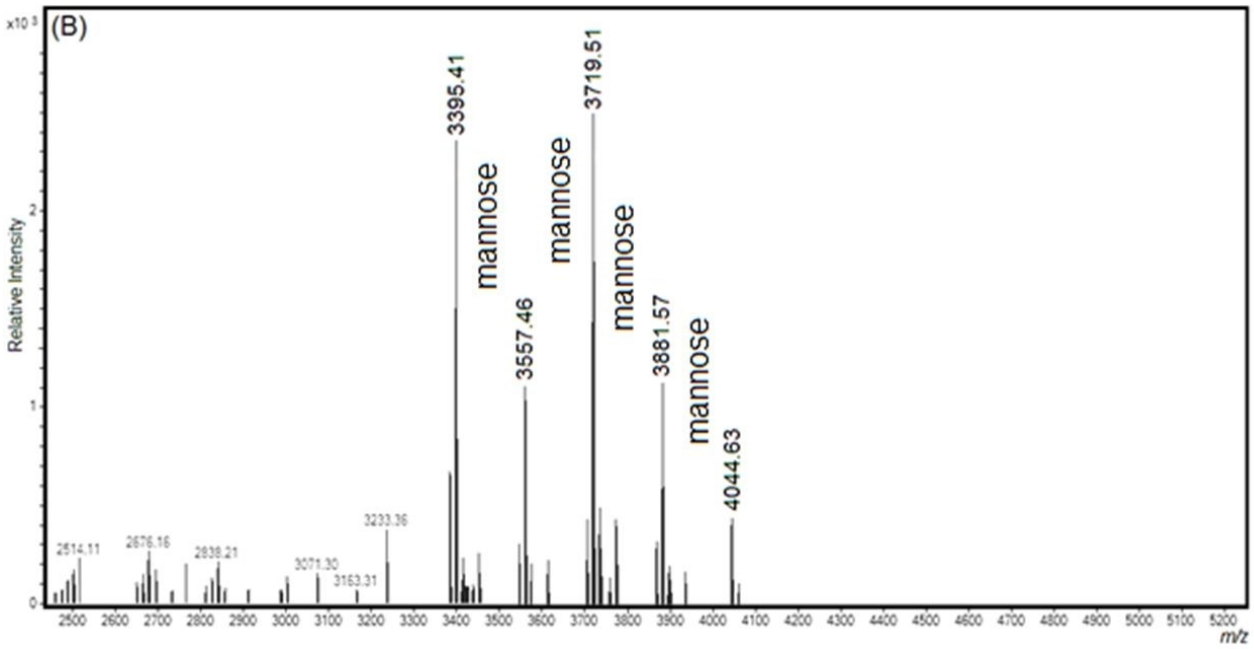


Figure 45: LCMS Fraction B

4.10 NMR data

Convincing NMR data on the purified mannan was difficult to obtain due to very small quantities of isolated material. However, within a mixture with the formerly characterized water-soluble *C. difficile* polysaccharides, previously undetected anomeric resonances could be observed using an HSQC experiment.

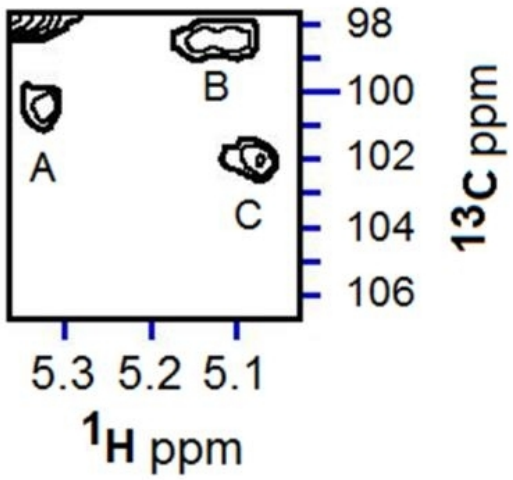
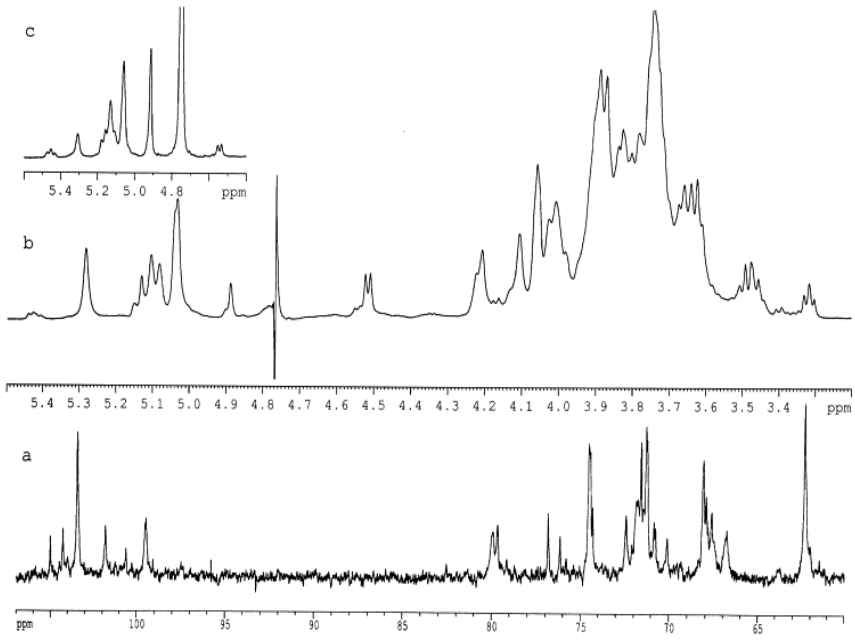


Figure 46: The anomeric region of the ^{1}H-^{13}C HSQC NMR spectrum of *C. difficile* water-soluble material, showing the anomeric resonances for the three main Man constituents. Anomeric signal A for 2-linked α-Man, signal B for 2/6-linked α-Man, and signal C for end-group α-Man.

The three characteristic anomeric resonances of α-Man units correspond to the same linkage types as the linkages observed from the previously completed linkage analysis. The chemical shifts of the anomeric protons were compared to previously published results from a different bacteria: *Pseudomonas syringae*.[45] These spectra are displayed in Figure 49. The anomeric signal A was found to be at δ_H 5.31/δ_C 100.5. This downfield proton peak allowed for the identification, when compared to the published data, of this cross peak as belonging to the 2-linked mannose linkage that was determined to be present via linkage analysis. The chemical shift of the B signal was determined to be at δ_H 5.11/δ_C 98.5. Comparing this to the published data led to the conclusion that these anomeric chemical shifts must correspond to a 2,6 linked mannose residue which confirms the finding of this linkage from the PMAA data. The chemical shift of the C signal was found to be at δ_H 5.08/δ_C 102.0. The downfield shift the carbon peak allowed for identification of the cross peak as one that belongs to a terminal mannose. These NMR findings confirmed the presence of all three linkages that were found from the linkage analysis.



Residue	H-1/C-1
1-P Man*p*	5.427 97.1
2-Man*p*	5.284 101.6
3-Man*p*	5.153 103.5
3-Man*p*	5.130 103.5
2,6-Man*p*	5.105 99.5
2,6-Man*p*	5.082 99.5
t-Man*p*	5.037 103.5
t-Man*p*	5.037 103.5

Figure 47: Proton and carbon spectra and assigned chemical shifts for the anomeric proton and carbon atoms from mannan fragments from *Pseudomonas syringae* [45]

4.11 Discussion

Collectively, the structural data that was obtained points to the presence of a branched mannan structure in the spores and in the cell wall of *C. difficile*. Chemical analysis revealed the presence of a terminal mannose acting as an end-group, 2-linked mannan as a linear structure in the cell wall and a 2,6-linked mannan as part of a branched structure. LC-MS analysis revealed that the extracted mannans were slightly heterogeneous in size with the molecular weights ranging around the mid 3000 Daltons. The NMR data validated the findings the GCMS and LCMS analysis and confirmed the presence of the three main mannan components, as anomeric resonances for the terminal α-Man, 2-linked α-Man and 2,6-linked α-Man were all observed. These observed resonances were consistent with reported signals for Man units in similar branched mannans in surface polysaccharides from other bacterial strains. [70]

Although they were detected in smaller amounts, the detection of other Man linkage types (3- and 6-linked) pointed to the presence of different structural regions within these Man units. A small number of 4-substituted glucose units were also detected, similar to those of previously characterized 4-linked glucan in other Clostridia species. [71] The observation that *C. difficile* spores contain glucans and now possibly mannans implies that carbohydrate-related genes must be present in the spores of *C. difficile*. These (1→4)-glucans with some 4,6-Glc branch points, have a structure akin to amylopectin. This glucan is expressed by animals and plants to serves a wide variety of functions. [72] This previous material from another strain that was shown to contain this 4-linked glucan, did not present mannan that was observed here, implying that these two surface polysaccharides are independent entities. Interestingly enough, it was noted that the yield of the mannan was proportional to the amount of spores present in the biomass. The mannans may be components of *C. difficile* spores as free oligosaccharides or

glycoprotein components. The mannans were confirmed from not being found in the growth medium of the bacteria by analyzing the growth medium by GC and GCMS.

Fungal, viral and bacterial species have been observed to produce branched mannan oligosaccharides and polysaccharides with similar structural units as the ones found in *C. difficile*. [45] In this work, it was observed that the detection of mannans occurred in *C. difficile* containing a high concentration of spore related material (five-day-old culture), and perhaps may play a role in the initial colonization of the bacteria in the gastro-intestinal tract. Possible glycosylation of spore-associated proteins have been postulated for *C. difficile* [73] and thus the mannan structures described here may be part of the glycoprotein framework in spores and/or vegetative cells. The wide distribution of this type of mannan across microbial species implies a key role in the pathogenesis of microorganisms, including that of *C. difficile*.

Chapter 5: Concluding Remarks

5.1 Conclusions

From GCMS and NMR analysis the structure of the CPS of HS: 44 was determined. It was determined that the CPS actually contains two CPS which is a rare occurrence when analyzing surface carbohydrate of *C. jejuni* . The first CPS was found to be identical in structure to a HS: 1 CPS containing galactose, fructose, O-methyl phosphoramidate and glycerol phosphate linking units. The second CPS was found to contain the 6-deoxy heptoses of galactose and altrose which are common heptoses found among various different serotypes of *C. jejuni* . However, separating the CPSs from one another did not go as planned and further structural analysis on the isolated HS: 44 heptose CPS could not be attempted. If separation was accomplished thorough structural analysis on the heptose containing capsule could be performed. However, the treatment method would also need to be sensitive to O-methyl phosphoramidate in an attempt to not cleave this crucial functional group from the CPS.

For *C. difficile*, it was concluded from the structural analysis that there are mannan units present in the cell wall of *C. difficile*. It was observed that a higher concentration of mannans was present in samples that contained more spore related material. These mannan structures are widely distributed amongst various microbial species. The presence of mannan sugars has been postulated to be various functions these polysaccharides can perform.. The mannans may be part of a glycoprotein framework in spores and or vegetatitve cells of *C. difficile* due to specific spore associated proteins that undergo glycosylation with mannose residues. Spores of *C. difficile* have also been found to contain 4-linked glucans similar in structure to the mannans that were observed. These glucans are used by various bacteria as crucial building blocks in various

biosynthetic processes used by the bacteria as well as playing key roles in host immune modulation. These glucans were also found in two other clostridium species *C. butyricum* and *C. botulinum*. In these species the glucans were used as carbon and energy sources for spore maturation. [74] These glucans were also theorized to be intracellular components that can be used by *C. difficile* as a resource for the synthesis of precursors required for the assembly of cell-wall carbohydrates or other biomolecules during germination processes of the bacteria. The glucans found on *C. difficile* were found to be required for the viability of the microbe. Because the glucans are required for the survival of the microbe they were found to be a possible target for curing and preventing *Clostridium difficile* infection. The function of the mannans found in *C. difficile* could possibly have the same function as the glucans and/or it was found that the mannan structures have similar importance to the microbe as the glucan structures the mannan structures found could potentially also be used as targets for curing and preventing *Clostridium difficile* infection.

5.2 Future Work

Future work on the HS: 44 CPS would be attempting various treatments and separation methods in an attempt to separate the HS: 44 capsule so that it can be analyzed in isolation from the HS: 1 capsule. This could be attempted with various acetic acid treatments that are used in the lab or more harsh basic treatments than the one that was performed. The products from the treatments could also be separated using column chromatography on top of the dialysis that was performed. Once the heptose containing capsule is isolated its structure would need to be determined. After finding the structure the goal would be conjugation to a carrier protein to create a carbohydrate vaccine and testing the serotype specific vaccine for its inclusion into a

single vaccine that can tackle the large number of diverse serotypes responsible for *C. jejuni* infection around the world.

Future work involving the mannan units would involve finding the exact functions of the mannan units for *C. difficile*. Figuring out the function of these mannans would go a long way in terms of determining if there are any potential medical applications for the mannans found in the *C. difficile* cell walls. Similar glucan structures were also found in *C. difficile* cell walls and were found to have crucial functions in the biosynthesis and energy production for *C. difficile* spores. Figuring out if the mannan structures are more important for the life of the spores or vegetative cells could also lead to interesting developments. The mannans were found more abundantly in the samples with higher spore concentration so it would be more logical to try to isolate larger amounts of these mannans for analysis from preparations of *C. difficile* spores. Once the function of these mannans is more understood they could possibly be used as targets for preventing and curing *Clostridium difficile* infection.

Chapter 6: References

1. Solomons, T.W.G.; Fryhle, C. B. *Organic Chemistry.* 10 th Ed: Wiley. Hoboken. 2011.

2. Stick, R. V.; Williams, S. *Carbohydrates: The Essential Molecules of Life.* 2 nd Ed: Elsevier Science. Amsterdam. 2010.

3. Robyt, J. F. *Essentials of Carbohydrate Chemistry*. Springer Verlag. New York. 1998.

4. Murray, R. K.; Granner, D. K.; Rodwell, V. W. *Harper's Illustrated Biochemistry*. 27 th Ed: McGraw–Hill. New York. 2006.

5. Bunn, H. F.; Higgins, P. J. *Science*, **1981**, *213*, 222-224.

6. Stryer, L. *Biochemistry*, 4 th Ed: W. H. Freeman and Company. New York. 1995.

7. Klemm, D.; Heublein, B.; Fink, H. P.; Bohn, A. *Angew Chem Int Ed*, **2005**, *44*, 3358-3393.

8. O'Sullivan, A. C. *Cellulose*, **1997**, *4*, 173-207.

9. Madigan, M. T.; Martinko, J.M.; Stahl, D. A. *Brock Biology of Microorganisms*. 13 th Ed: Benjamin-Cummings Publishing Company. San Francisco. 2012.

10. Bergey, D. H.; Holt, J. G.; Krieg, N. R.; Sneath, P. H.; Bergey's *Manual of Determinative Bacteriology.* 9th Ed: Lippincott Williams & Wilkins. Baltimore. 1994.

11. Lebeer, S.; Vanderleyden, J.; De Keersmaecker, S. C. J. *Nat Rev Microbiol*, **2010**, 8, 171-184.

12. Raetz, C. R. H.; Whitfield, C. *Annu Rev Biochem*, **2002**, *71*, 635-700.

13. Erridge, C.; Bennett-Guerrero, E.; Poxton, I. R. *Microbes Infect*, **2002**, *4*, 837-851.

14. Swoboda, J. G.; Campbell, J.; Meredith, T. C.; Walker, S. *ChemBioChem*, **2010**, *11*, 35-45.

15. Sorensen, M. C. H.; Van Alphen, L. B.; Fodor, C.; Crowley, S. M.; Christensen, B. B.; Szymanski, C. M.; Brondsted, L. *Front. Cell. Infect. Microbiol,* **2012**, *2*.

16. Davies, H. A.; Borriello, S. P. *Microb Pathogenesis*, **1990**, *9*, 141-146.

17. Nachamkin, I.; Blaser M. J. *Campylobacter*, 2nd Ed: ASM Press. Washington. 2000.

18. Miller, W. G.; Mandrell, R. E. *Campylobacter jejuni: New perspectives in molecular and cellular biology*. Horizon Scientific Press. Norfolk. 2004

19. Friedman, C. R.; Neimann, J.; Wegener, H. C.; Tauxe, R. V. *Campylobacter.* 2nd Ed: ASM Press. Washington. 2000.

20. Coker, A. O.; Isokpehi, R. D.; Thomas, B. N.; Amisu, K. O.; Obi, C. L. *Emerg Infect Dis*, **2002**, *8*, 237-243.

22. Jagusztyn-Krynicka, E. K.; Laniewski, P.; Wyszynska, A. *Expert Rev Vaccines*, **2009**, *8*, 625-645.

23. Ekdahl, K.; Andersson, Y. *Bmc Infect Dis*, **2004**, *4*, 54.

24. Advisory Committee on the Microbiological Safety of Food. *Advisory Committee on the Microbiological Safety of Food: Second Report on Campylobacter.* Food Standards Agency. London. 2005.

25. Wagner, J.; Jabbusch, M.; Eisenblatter, M.; Hahn, H.; Wendt, C.; Ignatius, R. *Antimicrob Agents Ch*, **2003**, *47*, 2358-2361.

26. Nachamkin, I.; Blaser, M. J.; Tompkins, L. S. *Camplylobacter jejuni: Current Status and Future Trends*. ASM Press. Washington. 1992.

27. Initiative for Vaccine Research. *State of the Art of Vaccine Research and Development.* World Health Organization. Geneva. 2006.

28. Gallardo, F.; Gascón, J.; Ruiz, J.; Corachan, M.; de Anta, M. T. J.; Vila, J. *J Travel Med.* **1998**, 5, 23–26.

29. McCarthy, N.; Giesecke, J. *Am J Epidemiol*, **2001**, *153*, 610-614.

30. Karlyshev, A. V.; Linton, D.; Gregson, N. A.; Lastovica, A. J.; Wren, B. W. *Mol Microbiol*. **2000**, *35*, 529-541.

31. Monteiro, M. A.; Baqar, S.; Hall, E. R.; Chen, Y. H.; Porter, C. K.; Bentzel, D. E.; Applebee, L.; Guerry, P. *Infect Immun*, **2009**, *77*, 1128-1136.

32. McNally, D. J.; Jarrell, H. C.; Li, J.; Khieu, N. H.; Vinogradov, E.; Szymanski, C. M.; Brisson, J. R. *Febs J*, **2005**, *272*, 4407-4422.

33. St Michael, F.; Szymanski, C. M.; Li, J. J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. *Eur J Biochem*, **2002**, *269*, 5119-5136.

34. Monteiro, M. A.; Baqar, S.; Hall, E. R.; Chen, Y. H.; Porter, C. K.; Bentzel, D. E.; Applebee, L.; Guerry, P. *Infect Immun*, **2009**, *77*, 1128-1136.

35. Chen, Y. H.; Poly, F.; Pakulski, Z.; Guerry, P.; Monteiro, M. A. *Carbohyd Res*, **2008**, *343*, 1034-1040.

36. McNally, D. J.; Jarrell, H. C.; Khieu, N. H.; Li, J.; Vinogradov, E.; Whitfield, D. M.; Szymanski, C. M.; Brisson, J. R. *FEBS J*, **2006**, *273*, 3975-3989.

37. Papp-Szabo, E.; Kanipes, M. I.; Guerry, P.; Monteiro, M. A. *Carbohyd Res*, **2005**, *340*, 2218-2221.

38. Hanniffy, O. M.; Shashkov, A. S.; Moran, A. P.; Prendergast, M. M.; Senchenkova, S. N.; Knirel, Y. A.; Savage, A. V. *Carbohyd Res*, **1999**, *319*, 124-132.

39. Gilbert, M.; Mandrell, R. E.; Parker, C. T.; Li, J. J.; Vinogradov, E. *ChemBioChem*, **2007**, *8*, 625-631.

40. Leuzzi, R.; Adamo, R.; Scarselli, M. *Hum Vacc Immunother*, **2014**, *10*, 1466-1477.

41. Ganeshapillai, J.; Vinogradov, E.; Rousseau, J.; Weese, J. S.; Monteiro, M. A. *Carbohyd Res*, **2008**, *343*, 703-710.

42. Knoop, F. C.; Owens, M.; Crocker, I. C. *Clin Microbiol Rev*, **1993**, *6*, 251-265.

43. Nazarko L. *Br. J. Community Nurs*, **2007***,* 12, 292–295.

44. Wilkins, T. D.; Lyerly, D. M. *J Clin Microbiol*, **2003**, *41*, 531-534.

45. Kyne, L.; Hamel, M. B.; Polavaram, R.; Kelly, C. N. P. *Clin Infect Dis*, **2002**, *34*, 346-353

46. Pepin, J.; Valiquette, L.; Alary, M. E.; Villemure, P.; Pelletier, A.; Forget, K.; Pepin, K.; Chouinard, D. *Can Med Assoc J*, **2004**, *171*, 466-472.

47. Pepin, J.; Valiquette, L.; Cossette, B. *Can Med Assoc J*, **2005**, *173*, 1037-1042.

48. McCallum, M.; Shaw, S. D.; Shaw, G. S.; Creuzenet, C. *J Biol Chem*, **2012**, *287*, 29776-29788.

49. Wong, A.; Lange, D.; Houle, S.; Arbatsky, N. P.; Valvano, M. A.; Knirel, Y. A.; Dozois, C. M.; Creuzenet, C. *Mol. Microbiol*, **2015**, *96*, 1136-1158.

50. Stern, A. M.; Markel, H. *Health Affair*, **2005**, *24*, 611-621.

51. Hansen, B. *Am Hist Rev*, **1998**, *103*, 373-418.

52. Atkinson, W. L.; Pickering, L. K.; Schwartz, B.; Weniger, B. G.; Iskander, J. K.; Watson, J. C. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control*, **2002**, *51*, 1-35.

53. Dranoff, G. *Nat Rev Cancer*, **2004**, *4*, 11-22.

54. Macleod, C. M.; Hodges, R. G.; Heidelberger, M.; Bernhard, W. G. *J Exp Med*, **1945**, *82*, 445-465.

55. Austrian, R.; Douglas, R. M.; Schiffman, G.; Coetzee, A. M.; Koornhof, H. J.; Haydensmith, S.; Reid, R. D. W. *Clin Res*, **1976**, *24*, A483-A483.

56. Goldblatt, D. *Clin Exp Immunol*, **2000**, *119*, 1-3.

57. Astronomo, R. D.; Burton, D. R. *Nat Rev Drug Discov*, **2010**, *9*, 308-324.

58. Pollard, A. J.; Perrett, K. P.; Beverley, P. C. *Nat Rev Immunol*, **2009**, *9*, 212-220.

59. Guerry, P.; Poly, F.; Riddle, M.; Maue, A. C.; Chen, Y.-H.; Monteiro, M. A. *Front. Cell. Infect. Microbiol*, **2012**, *2*, 7.

60. Monteiro, M. A.; Baqar, S.; Hall, E. R.; Chen, Y. H.; Porter, C. K.; Bentzel, D. E.; Applebee, L.; Guerry, P. *Infect. Immun,* **2009**, 77, 1128–1136.

61. Monteiro, M. A.; Ma, Z. C.; Bertolo, L.; Jiao, Y. N.; Arroyo, L.; Hodgins, D.; Mallozzi, M.; Vedantam, G.; Sagermann, M.; Sundsmo, J.; Chow, H. *Expert Rev Vaccines*, **2013**, *12*, 421-431.

62. Chow, H.; Sagermann, M.; Ma, Z.; Vandentam, G.; Monteiro, M. A. Vaccination with Clostridium difficile PSII polysaccharide antigen adjuvanted with KLH induced broad-based enhancement of adaptive immune responses and protection in mice. 7th Vaccine and ISV Congress. Sitges, Spain. October 27-29, 2013; and Pequegnat, B.; Sagermann, M.; Arroyo, L.; Chow, H.; Monteiro, M. A. An anti-C. difficile PSII polysaccharide-KLH conjugate vaccine is efficacious in mice. International Conference on the Molecular Biology and Pathogenesis of the Clostridia (ClostPath 8) Queensland, Australia, October 22-26, 2013.

63. Westphal, O.; Jann K. *Methods Carbohydr Chem.* **1965**. 5, 83-91.

64. Skoog, D. A.; West, D. M.; Holler, F. J. *Fundamentals of Analytical Chemistry*, 9 th ed: Brooks-Cole. Belmont. 2013.

65. Biermann CJ, McGinnis GD. *Analysis of Carbohydrates by GLC and MS*, CRC Press, Inc. Boca Raton. 1989.

66. Friebolin, H. *Basic One and Two-Dimensional NMR Spectroscopy*, 5th Ed.: Wiley-VCH. Weinheim. 2011.

67. Agrawal, P. K. *Phytochemistry*, **1992**, *31*, 3307-3330.

68. Gregoriadis G, Division NATOSA. New generation vaccines: the role of basic immunology ; [proceedings of a NATO Advanced Study Institute on New Generation Vaccines: The Role of Basic Immunology, held June 24 - July 5, 1992, in Cape Sounion Beach, Greece]: Springer London, Limited, 1993.

69. Bax, A.; Summers M. F. *J. Am. Chem. Soc*, **1986,** 108, 2093.

70. Corsaro, M. M.; Evidente, A.; Lanzetta, R.; Lavermicocca, P.; Molinaro, A. *Carbohyd Res*, **2001**, *330*, 271-277.

71. Bertolo, L.; Boncheff, A. G.; Ma, Z. C.; Chen, Y. H.; Wakeford, T.; Friendship, R. M.; Rosseau, J.; Weese, J. S.; Chu, M.; Mallozzi, M.; Vedantam, G.; Monteiro, M. A. *Carbohyd Res*, **2012**, *354*, 79-86.

72. Gilbert, R. G. *Anal. Bioanal. Chem*, **2011**, *399*, 1425-1438.

73. Dingle, K. E.; Didelot, X.; Ansari, M. A.; Eyre, D. W.; Vaughan, A.; Griffiths, D.; Ip, C. L. C.; Batty, E. M.; Golubchik, T.; Bowden, R.; Jolley, K. A.; Hood, D. W.; Fawley, W. N.; Walker, A. S.; Peto, T. E.; Wilcox, M. H.; Crook, D. W. *J Infect Dis*, **2013**, *207*, 675-686.

74. Strasdine, G. A. *Can. J. Microbiol*, **1972**, 18, 211-217.